

1 AAGCGATAGC TGAGTGC CGGC GGCTGCTGAT TGTGTTCTAG GGGACGGAGT
 51 AGGGGAAGAC GTTGCTCTC CCGGAACAGC CTATCTCATT CCTTTCTTTC
 101 GATTACCGT GGC GCG GAGA GTCAGGGCGG CGGCTGCGGC AGCAAGGGCG
 151 GCGGTGGCGG CGGC GGG CAGC TGCAGTGACA TGTCCAGCAT GAATCCCAGA
 201 TATGATTATT TATTCAAGTT ACTTCTGATT GGC GACTCAG GGGTTGGAAA
 251 GTCTTGCCTT CTTCTTAGGT TTGCAGATGA TACATATACA GAAAGCTACA
 301 TCAGCACAAAT TGGTGTGGAT TTCAAAATAA GAACTATAGA GTTAGACGGG
 351 AAAACAAATCA AGCTTCAAAT AGAGTCCTTC AATAATGTTAACAGTGGCT
 401 GCAGGAAATA GATCGTTATG CCAGTGAAAAA TGTCAACAAA TTGTTGGTAG
 451 GGAACAAATG TGATCTGACC ACAAAAGAAAG TAGTAGACTA CACAACAGCG
 501 AAGGAATTG CTGATTCCCT TGGATTCCG TTTTTGGAAA CCAGTGCTAA
 551 GAATGCAACG AATGTAGAAC AGTCTTCAT GACGATGGCA GCTGAGATTA
 601 AAAAGCGAAT GGGTCCCAGA GCAACAGCTG GTGGTGCTGA GAAGTCCAAT
 651 GTTAAAATTC AGAGCACTCC AGTCAAGCAG TCAGGTGGAG GTTGCTGCTA
 701 AAATTGCGCT CCATCCTTT CTCACAGCAA TGAATTGCA ATCTGAACCC
 751 AAGTGGAAAAA ACAAAATTGC CTGAATTGTA CTGTATGTAG CTGCACTACA
 801 ACAGATTCTT ACCGTCTCCA CAAAGGTCAAG AGATTGTAAGA TGGTCAATAC
 851 TGACTTTTTT TTTATTCCCT TGACTCAAGA CAGCTAACTT CATTTCAGA
 901 ACTGTTTAA ACCTTTGTGT GCTGGTTTAT AAAATAATGT GTGTAATCCT
 951 TGTTGCTTTC CTGATACCAG ACTGTTCCC GTGGTTGGTT AGAATATATT
 1001 TTGTTTGAT GTTATATTG GCATGTTAG ATGTCAGGTT TAGTCTTCTG
 1051 AAGATGAAGT TCAGCCATT TGTTCAAAAC AGCACAAGCA GTGTCTGTC
 1101 CTTTCCATGC ATAAAGTTA GTGAGATGTT ATATGTAAGA TCTGATTGCG
 1151 TAGTTCTTCC TTGTTAGAGTT ATAAATGGAA AGATTACACT ATCTGATTAA
 1201 TAGTTCTTC ATACTCTGCA TATAATTGTT GGCTGCAGAA TATTGTAATT
 1251 TGTTGCACAC TATGTAACAA AACAACTGAA GATATGTTA ATAAATATTG
 1301 TACTTATTGG AAGTAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA
 1351 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA
 1401 AAAAAA (SEQ ID NO:1)

FEATURES:

5'UTR: 1-179
 Start Codon: 180
 Stop Codon: 699
 3'UTR: 702

Homologous proteins:

Top 10 BLAST Hits

	Score	E
CRA 108000024647144 /altid=gi 12728868 /def=ref XP_002675.2 RA...	372	e-102
CRA 18000004923424 /altid=gi 4758988 /def=ref NP_004152.1 RAB1...	332	5e-90
CRA 18000004937406 /altid=gi 131787 /def=sp P05711 RAB1A_RAT RAS...	328	1e-88
CRA 18000004952860 /altid=gi 131785 /def=sp P22125 RAB1_DISOM R...	320	3e-86
CRA 18000004995539 /altid=gi 103720 /def=pir D38625 GTP-bindin...	313	3e-84
CRA 18000004967528 /altid=gi 92339 /def=pir S06147 GTP-binding...	297	2e-79
CRA 18000004880958 /altid=gi 464524 /def=sp Q05974 RAB1_LYMST R...	282	9e-75
CRA 18000004908714 /altid=gi 466171 /def=sp P33723 YPT1_NEUCR G...	253	3e-66
CRA 18000005175724 /altid=gi 7497231 /def=pir T33781 hypothetical...	253	4e-66
CRA 335001098696672 /altid=gi 11558649 /def=emb CAC17833.1 (AJ...	251	2e-65

BLAST dbEST hits:

	Score	E
gi 12867866 /dataset=dbest /taxon=960...	654	0.0
gi 12097820 /dataset=dbest /taxon=96...	654	0.0
gi 12793758 /dataset=dbest /taxon=960...	624	e-177
gi 12338056 /dataset=dbest /taxon=96...	622	e-176
gi 11977068 /dataset=dbest /taxon=96...	609	e-172
gi 10339840 /dataset=dbest /taxon=960...	517	e-145
gi 10349761 /dataset=dbest /taxon=960...	436	e-120
gi 10997958 /dataset=dbest /taxon=96...	385	e-105
gi 10996533 /dataset=dbest /taxon=96...	381	e-103

EXPRESSION INFORMATION FOR MODULATORY USE:

library source:

From BLAST dbEST hits:

gi|12867866 Fetal brain
gi|12097820 Adrenal gland
gi|12793758 Brain neoroblastoma cell line
gi|12338056 Adrenal gland
gi|11977068 Skin melanotic melanoma
gi|10339840 Uterus leiomyosarcoma
gi|10349761 Skin melanotic melanoma
gi|10997958 Placenta
gi|10996533 Placenta

From tissue screening panels:

Whole brain

1 MSSMNPEYDY LFKLLLIGDS GVGKSCLLR FADDTYTESY ISTIGVDFKI
51 RTIELDGKTI KLQIESFNNV KQWLQEIDRY ASENVNKLKV GNKCDLTTKK
101 VVDYTTAKEF ADSLGIPFLE TSAKNATNVE QSFMTMAAEI KKRMGPGATA
151 GGAEKSNVKI QSTPVKQSGG GCC (SEQ ID NO:2)

FEATURES:

Functional domains and key regions:

[1] PDOC00001 PS00001 ASN_GLYCOSYLATION
N-glycosylation site

125-128 NATN

[2] PDOC00005 PS00005 PKC_PHOSPHO_SITE
Protein kinase C phosphorylation site

Number of matches: 5

1	59-61	TIK
2	97-99	TTK
3	98-100	TKK
4	106-108	TAK
5	122-124	SAK

[3] PDOC00006 PS00006 CK2_PHOSPHO_SITE
Casein kinase II phosphorylation site

Number of matches: 3

1	35-38	TYTE
2	106-109	TAKE
3	127-130	TNVE

[4] PDOC00007 PS00007 TYR_PHOSPHO_SITE
Tyrosine kinase phosphorylation site

30-36 RFADDTY

[5] PDOC00008 PS00008 MYRISTYL
N-myristoylation site

Number of matches: 3

1	21-26	GVGKSC
2	147-152	GATAGG
3	152-157	GAEKSN

[6] PDOC00017 PS00017 ATP_GTP_A
ATP/GTP-binding site motif A (P-loop)

18-25 GDGGVGKS

[7] PDOC00579 PS00675 SIGMA54_INTERACT_1
Sigma-54 interaction domain ATP-binding region A signature

14-27 LLLIGDSVGKSQL

FIGURE 2, page 1 of 2

BLAST Alignment to Top Hit:

>CRA|108000024647144 /altid=gi|12728868 /def=ref|XP_002675.2| RAB1,
member RAS oncogene family [Homo sapiens] /org=Homo
sapiens /taxon=9606 /dataset=nraa /length=222
Length = 222

Score = 372 bits (944), Expect = e-102
Identities = 190/222 (85%), Positives = 190/222 (85%), Gaps = 32/222 (14%)
Frame = +3

Query: 129 GGCGSKGGGGGGSCSDMSSMNPEYDYLFKLLLIGDSGVGKSCLLRFADDTYTESYIST 308
GGCGSKGGGGGGSCSDMSSMNPEYDYLFKLLLIGDSGVGKSCLLRFADDTYTESYIST
Sbjct: 1 GGCGSKGGGGGGSCSDMSSMNPEYDYLFKLLLIGDSGVGKSCLLRFADDTYTESYIST 60

Query: 309 IGVDFKIRTLGDGKTIKLQI-----ESFNNVK 392
IGVDFKIRTLGDGKTIKLQI ESFNNVK
Sbjct: 61 IGVDFKIRTLGDGKTIKLQIWNDTAGQERFRTITSSYYRGAHGIIVVYDVTQESFNNVK 120

Query: 393 QWLQEIDRYASENVNKLLVGNKCDLTTKVVDTTAKEFADSLGIPFLETSAKNATNVEQ 572
QWLQEIDRYASENVNKLLVGNKCDLTTKVVDTTAKEFADSLGIPFLETSAKNATNVEQ
Sbjct: 121 QWLQEIDRYASENVNKLLVGNKCDLTTKVVDTTAKEFADSLGIPFLETSAKNATNVEQ 180

Query: 573 SFMTMAAEIKKRMGPAGATAGGAEKSNSVQIQSTPVKQSGGGCC 698
SFMTMAAEIKKRMGPAGATAGGAEKSNSVQIQSTPVKQSGGGCC
Sbjct: 181 SFMTMAAEIKKRMGPAGATAGGAEKSNSVQIQSTPVKQSGGGCC 222 (SEQ ID NO:4)

Hmmer search results (Pfam):

Model	Description	Score	E-value	N
PF00071	Ras family	256.4	7.7e-75	2
CE00060	CE00060 rab_ras_like	170.0	3.9e-47	2
PF00634	BRCA2 repeat.	9.9	0.39	1
PF00056	lactate/malate dehydrogenase	3.9	3.4	1

Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
PF00056	1/1	13	29 ..	1	18 [. .]	3.9	3.4
CE00060	1/2	8	64 ..	20	77 ..	86.8	8.9e-23
PF00071	1/2	13	64 ..	1	52 [. .]	111.9	4.8e-32
PF00634	1/1	57	79 ..	13	35 .]	9.9	0.39
CE00060	2/2	65	140 ..	110	188 ..	81.2	2.9e-21
PF00071	2/2	65	173 .]	85	198 .]	142.4	4.5e-41

1 TTTTGGGTGT GTGTGTGT GTGTGTGT GTGCCTTAC TAGTGACTCA
51 GGTACAGTT TTCTGAGATT TTTTTCTCC CCTCAAGACA GAATCTTGCT
101 CTGTCGCCA GGCTGGAGTG CAGTGGCCTC TCGGCCCACT GTAGCCTCCG
151 CCTCCCGGGT TCAAGCAATT TTCCTGCCTC AGCCTCCCAG GTAGCTGGGA
201 TTACAGGCAC GCGCCACCAT GCCTGGCTAA TTTTTGTATT TTTAGTAGAG
251 ACAGTGTTC ACCATGTTGG CCAGGCTGGT CTTGAATTCC TGACCTCGTG
301 ATCTGTCCGT TTTGGCCTCT CAAATTCTCG AGATTACAGG CATGAGCCAC
351 CGAGCCTGGC CAGTTTCTG AGTTTTTATT TGAAATCAA ATAAGCTTT
401 TTTTTTTTT TAATGGGCTT TAGAGTCCAG GGTAAACGAAC ACTTTTGGT
451 GCCTATTACT GAACCATTCA GGGTATTCCG GGGGTGGTGA CCGTGTTCAT
501 TTCAGAAACC AACATGTTCA TTTCAGAAAC CAAACTCGGG TAACTTTGA
551 TAAGTTCATC AACTAAGGCC CATGGCAGAA TTTGAGGGCT AAGGGGTGTA
601 ATTAGTGTAT GGGTAGAAAT AAGTGCCTTC TTTCTATATT TTGGCGTTGT
651 AGGAATTAA AGTGATTCTG CAGTAAGTCT CAGGAGACAA TTTCTTAGT
701 TCTTAAAGT TGGAAGATAA ACTTTGGACA ATGTATTACA CTATGCCCTT
751 TGTAATTAAA TAACTCAAGA TAATGTGTTA AAGTTAGCG GAGATTTAAA
801 TTCTGAGCT GATTAAAGAG AGCTGTTAAG GCCATAGGTT TTTAAAAAAT
851 GAGTTAATAT TACTCCCAGA AATTGTAGGC ACTATATAGT GATGAATTGC
901 ATATTTTAT TGCTTATTAT TTTCCAGTCT TGCAGAATGG CTCAGGGTTA
951 GTAGCAACTA AAAGATAATA CATTACAATT CAACCTGAAG GCCGGGACGA
1001 AGGTAGGAAT TGGATTTAG GCTGGCTCTG GGCTGTTGTC CTCATGCCA
1051 TGGGATGTGG AGCCATTGAA GGTTGTTGGG TCACGATGCA GGTGCTGTCT
1101 CAGAAAGATA CATCCGACTG TGTGTCAAA TGGGCTGGGG CGGAGAAGAG
1151 AGAGAGAGGT AGAGTCCATT TGGAGACTAC TGCAATAGCC AGGCTGACGA
1201 GTTAAGAGCG GGGCACAGTA AGAATGGAA GAAATCTAAG AAGAAAATGG
1251 TAGTGCACGG GGCCAACAAT GGACGATGAC CGAACCCAGG TGGGGATGGG
1301 TGAGTACGA GAAGAACCGC TCCGTGCCGT CCAGGGAGCC CCTTGACTTC
1351 CCTTCTGTT TTAGAGCGGA CGTCCTCCTA CCAGCCCCCA ACCAGCGCCA
1401 CCAGGGTGGC GCAAGCCTCA AGCTGGTCAG GTCAGCAACA GCCGCAACGG
1451 AGGCAGGAGC CGACACGCTC GTACCCCGGC CCCCTCCCCG CCCCCGCACC
1501 CCCGCAGTC CCTCCGGTTT GACCACTCCC CCCGGTCCCT TGCCCTCCCC
1551 GACCCCCCAGC CCTCCGTCGGC CGCCGGCACC ACCCTCCGCC CCTCTCCGCC
1601 CCCTCCCCCG TGGGCGCTG ACTCGCCCGG CTGCCACGTC TCACTGATGA
1651 CATCACTAGG GCAGCTCGGC TTAGCCAAT CCGCCAGGGG GAGTCCGAGC
1701 GAAGTCCTAG CCAGCGAGTC AGAGGGGAGG GGAGCAGGGG GGGCCGAGG
1751 GTGGGGAGGT GAGGGAGTGG GGAATGGGGC GGGCGACAAC CCTTCAGGTA
1801 CGCATGCCCG AGAGGCGCGG CGCTTGGCGG GAAGCTGAGT CCTGGCCTTG
1851 CGTCGCACTG TCTGCTCTCA GCTCGCGTAG CCGCGCTCGC GACTCCCTT
1901 CCCGCATGC CAGCGGTGC GGCGGCCCTC TGGGCGTGT AAAGGCCCC
1951 CGGTCTAAGG CCTCCCTATT TCCTGGTTCG CCGCGGCCA TTTTGGGTGG
2001 AAGCGATAGC TGAGTGGCGG CGGCTGCTGA TTGTGTTCTA GGGGACGGAG
2051 TAGGGGAAGA CGTTTGTCT CCCGGAACAG CCTATCTCAT TCCTTCTTT
2101 CGATTACCCG TGGCGCGGAG AGTCAGGGCG GCGGCTGCAG CAGCAAGGGC
2151 GGCAGTGGCG CGGGCGCAG CTGCACTGAC ATGTCCAGCA TGAATCCGA
2201 ATAGTGAAGT CAGGAGAGCA CCGGTCGGCT GGGTCGGTGG GCCAGCTTGG
2251 GGGATCTTAA AGGGGTCAGAG GAGGGTTGGG GCAGAACTCG GGGCATCGGC
2301 TGGGGTGGAG CGAGGGTGTAG GGTCAGGAG AGGCTGGCGG CGGGGAGTCG
2351 GGGCCCATG TCTGACCGGG AGGGCCGCC GCGCGGGGAG GGGTCGGGC
2401 CGGAGGGGTG AGCCGCCCGG CCCTGGACCG GGTCAAGGTTA GAGGGCCTGA
2451 CTGCGGGCGG GGTGCTGAGG AAGCCTGCCG AGGGGCTGG GCGGGTGTGA
2501 AGGGTATCT TCTCTCGGAG GCACTGACTT TTGAAGGAGG ACTTGTCT
2551 AAGGGGAGGG GATGGGGTGG GAGAGCCCTT CTAGAGGCCA CTGTCAGACC
2601 CTGCGCCCGC ACTCTGCGGA GCTGTCAGGA TCTTCGGGT AGAAACCGC
2651 TTACTTGTAA AATCTGAGC TTGTTGGGT TCTCTCCCTC CATCCTCCCC
2701 GCCAGGTTTC AGGTAATATG GATGCTTTT GGGACTGCGT GGGATTGAGG
2751 GGAATGAGTA GATGGTGAGA AGCAACTGAA CATTATATTAG TTCTCTTTT
2801 GAGTTGTGTC TTGGAGGAGT TGTGAGGAG CTCGCCGGGT CCATTGCCCT
2851 CCTATAAAA CCTGGCATT TGTGAGAATT TTGTTTTTT TTTTTTAA
2901 GAGGACACCT AAGTCATTT GTCTTCTGTG GGTCAAGGG AAAA
2951 ACTAAAGCCA AGAAATGTCT TTTTGATACT CGCAGATTA AGGAAGCTTG
3001 CTGTCAGTT GAAAGAGAAA CGAACGGGAC CTATGATAGA TCTGTATGTA
3051 GGTGGGAT TACCTGCTTG GATGCTTGCA GATAGGGAAAT GAGGGTCCAT
3101 GACGTGTCAT GAAAAGTTAA TGCATTCTT TTTCTGCTT ACTCAAGAAG

FIGURE 3, page 1 of 21

3151 TCACCACAGC AGATGTGACA CACCTGGCAC CTTTCCTGGG AACTGGTGT
3201 CACTCCCTT GGGTAGAGTT TGTTGGGCTC TCCTCAATGG CCCTTTAAAA
3251 ATTCCTCTA CAGTTTACAT GCATGTAAAG TAATGAATAA TTGGAAGAGA
3301 CGAATTGGT ATTCTTTTC AGTGTCAAAG GCCTTGAGG GATGGGGGAA
3351 AATCAGTATT TGTTGTAAA GTTGAGTTA TTTGCTGGTT TGGTCAATT
3401 CTGCTAGACA TTTCCCCTA AAAGGTCCAC CCACCAAGTT AGCTGACTGT
3451 CATATGTGTG TCACATGGCT CTTGCAAAAT GCTTACAAGT TTTGTAATAG
3501 TGTGGCTTGA AGCTGAAATC TTTGCACTA AACAGAAACC GTAGTATTT
3551 ATTAGAATT CATGCTTAG AAGTTGAGGG TAGTGTCTT GTAGTGCACAT
3601 TTGCTGTGTT GACAGTTAA AAAAATTTT TTTTCAAGGG CTCCAAGGAC
3651 AAAGTTGGTT TTGACAGTT GAACGGAGGT GAACTTGAGG TTCTTAATTT
3701 AGTAGTTTC TTGGTAACAA TAAAGAACAT GGATTTACTG CTTTATCGAG
3751 GTTTATAGAC CTCTACTGTT CAGGAAATT TCTGAATTG CTATATATAT
3801 GTTTATTAGT GTAATAAAAT CTTCAAGATT AGTTGAGAAC TTGACAAGT
3851 TACTCAGCCT CTGAATT TTTCCCTTT GTAAAATAGG ATAATTGGAG
3901 TCATTATTCC TGTCAAGGGTA GTGGTAAAT TCAAATGTAT ATAAAAGAAT
3951 TTGAAAACACT GTGTGAGCAT TCTTCAGGTG GTATGCATCA TTTCATGAA
4001 AGGCATTCTA TTAGTACCAAG GATTAGGAA TATAATCCTT GCGCTTAAGA
4051 AGTTTAGATA TAGGCCAGGC GCGGTGGCTC ACCTCAGTAA TCCCAGCACT
4101 TTGGGAGGCC GAGGGGGCG GATCCCGAGG TCAGGAGATC GAGACCATCC
4151 TCGGTAACAC GGTGAAACCC CGTCTCTACT AAAAATGCAA AAAAATTAGC
4201 CGGGCGTGGT GGTGGCACC TGTAGTCCCC GCTACTCGAG AGGCTGAGGC
4251 AGGAGAATGG CGTGATCCCC GGAGGTGGAG CTTGCAGTGA ACCAAGATCT
4301 GGCCACTGCA CTCCAGCCTG GACGACAGAG CAAGACTCCG TCTCAAAAAA
4351 AAAATTATT ATTGTTTGA GACGGAGTTT CAATCTGTT GCCCAGGCTG
4401 GAGTGCAATG GCGCAAATCT CCTCTCACCG CCACCTCCGC CTCCTGGGTT
4451 CAAGTGATTC TCCAGCCTCA GATTCCCCAGA AAGTTGGGAT TACAGGCATG
4501 TGCCACCACT CCCGCTAAT TTTGTTTTTG TGGTAGAGAC GGGGTTCTC
4551 CATGTTGGTC AGGCTGGTCT CAAACTCCCG AAGTGTACCG CCGGCCTCAG
4601 CTTCCCAAAG TGTTGGGATT ACAGGCGTGA GCCACCGCGC CCGGCAGAAA
4651 TAGATTTTAT ACATGTCAAA TACCAAGTACA TATAGCAAAT TCCAGATGTG
4701 TGGCATGGAT GAGAGCAACA AGATTTCAAGG GGGATGGTGG GTTGTGGTTG
4751 GCTATCTGGG TTTTGGAAAGA CTTTATAGAA GAGAGACCTG AAAGGGATT
4801 ATCAGCAATT AGATTTGGAG GAACAGAGGG AGTGAATAGG AATTTCAAG
4851 GGGGAGAAGA AGGAGGAATG GCTCATAAAT GACAAGGACA GTATAAAGTA
4901 AATACGGTGT CAAATCATCC TTTCTTTGA AGACTAATGA CCTCAAAGGG
4951 ATCAAACCCA GAAACAGTTT TTATATTTTT TCTGGGATCA AATACATGGG
5001 TATCTGGCT ACTATATTG TATTCTAGAC TGTTAGTAA AATAATACAG
5051 GAATTTGAGA AAACCTTGC AAAAGTGTAA GTGAAAATTA CTTAGGGTGA
5101 GAGGAAGTGA GGGATTTTT ATTAGGGAG GTCACAAGGG CAGTGAGCAA
5151 TCAGATTTT AGTAATCTGA CTTAACAGT TTCTTTTGT TTTAATGAAG
5201 CTTGTTATCT TTATAAAAGT AATTAGAGAA AATTTGAGAA ATAAAGGAAA
5251 GAAAGAAAAG TTCTTAGTG TTTTATCAGC CAAATACAAG CTCATTGTT
5301 TTTAACATCT TGTTCCAAAC TCCAAAGTCT TGCTTCTCT TCAATTAAAA
5351 CTTTAATGGG TGGATGCTT TCCTGCTTCC AGTATGTTAT CTTAATAACT
5401 AACAAATGGTA TATTAGCTAA TGTTTACAAA TGTAATCCAG ATGTTCTTA
5451 AGTTACTTTG GTTTATCATT ACCAATTAT ATTGTTCTT TTAGAAATTT
5501 ATAATCTTG TTAATGGGTT CTGCTAAATT TGTTAGTGAA AATGGGATCT
5551 TGAGAAAAAA GATTCTGAAG CAACAGAAATT TTGAGATTAA TATTGGTTTA
5601 CATAAGAGTT GGTAGCTGTA TTACTTTTT TGTTTGTGTT GTTTTTTTTT
5651 TGAGACGGAA TCTTGCTCTG TCGCCCAAGGC CTTGGCCTCC CAAAGTGTG
5701 GGATTACAGG CGTGAGCCAC TGTCCTGGC TGTTGTGTT TTTTTTTGTT
5751 TTTGTTTCT TTTCTTTTCA TTGTTTCA GATGGAGTCT CACTCTGTCA
5801 CCCAGGGCTGG AGTGCAGTGG CGCGATCTTG GCTCACTGCA ATCTCTGCCT
5851 CCTGGGTCTA AGCGATTTTC CTGCTTGGT CTCCTGAGTA GCTGGGATTA
5901 CAGGCATTG CCACCAATAAC CAGCTAAATT TTGTTAGAG TACCCAGCCA
5951 TCTCTAATGT TGATCAGGCT GAAGCAGGTG GATCACCTAA GGTCAAGGAGT
6001 TCAAGACCCAG CCTGGCAAT ATGGCAAAAC CCTATCTCTA CTAATACAGA
6051 AAATTATCTG GGTGTGTTGG CTGGCGCTGG TAATCCAGC TACTCGGGAG
6101 GCTGAGGCAAG GACAATCTCT TGAACCTCGG AGGTGGAGGT TGCAGTGAGC
6151 CGAGATCACA CCATTGCACT CCAGCCTGGG CAACAGAGCA AGACTTGTCT
6201 CAAAAAAA AAAAAAAA AAAAAAAGGC AATTGAAAGT GTAATCTGAA
6251 CAGTTAAAAA AGTAGATAGA AAGGGTTAAA GCTTTTTT GAGGATCTGA

FIGURE 3, page 2 of 21

6301 AGAAAAATGT GGATTTTT TGAGCTACGT TTTGAAGCAG GCAGTGATTA
 6351 TTTCAGCACA TTAAGAAATG CTTAACATGG CCAGGCCAG TGGCTCACGC
 6401 CTGTAATTCT CAGCACTTGG GAGGCCAG GTGGCCGGAT CATTGAGGT
 6451 CATGACCAGC CTGGCCAACA TGATGAGACA CTGCCTCTAC TAAAAAATACA
 6501 AAAATTAGCT GGGTGTGGTG GTGCACGCCT GTAATTCCAG CTACTCAGGA
 6551 ACCTGAGGCA GGAGAGTCAC TTGAACCTGG GAGGCCAGG CTGCAGTGAG
 6601 TCCAGATCAT GCCACTGCAC TCCAGCCTGA GGGACAGAGT GAGACTCCTC
 6651 AAAAAAAA AAAAAAAAG AAAGAAATAC TTAACATTAT TCTCGTGATT
 6701 ATTCTCATAA CATTTCAT AATCCACTGG CTTCCAGTGG ATTTTTTAG
 6751 TGTCAAGAAA ATAATTTGA TTGGTTCATC TTTAAGGAAT GTGTTAAGAA
 6801 TAAAGCATGT CTACCTGTCT TCAGTATAACC AGCTAACTAT AGTAGGAAGA
 6851 AATATAGTAG TCTACTTAGA TCAACTATAA TTCTTTAATG CAGAAAAAGT
 6901 TTAAAGTATT TACCTTATT TTAGCCCCCA TCCCCTTAAG TATATCATGG
 6951 CTCCAGAATC TCTGAAAATG TTATCAGTCT TTCAGACTTT GCTCTCTT
 7001 CATGTTATAC TCAAGAAACA TTGACCTTT TTTTTTTTT TTTGCTTGC
 7051 ATTGTGTTTC AAATAATTAA TAACAAAACT TAAAGTGGTAAAG
 7101 CAGGTTGTCT TTGTAACCTT TGGTGGTGGT TTGAAAAACT CAGAAAAAGTT
 7151 TAAAGAAGAA AGATAACTAG TATTCTCATT GTCCAGAATA TGATTTTTA
 7201 AATGTCTATA GAATATCACC ATCTGTAATT CTTCCGGTAA TTTAAGTATT
 7251 CAGTAGTTGT ATAAAACCTT TAAATATAT ATATTGAGAA TTTTGTGTGA
 7301 ATGAGATGAT GAGATAATCT TGTAGGATCA TTTAAAGATA AGAACTGAGG
 7351 CCTGGCACAG TGGCTCATGC CTATAATCAC AGCACTTGG GAGGCCAGG
 7401 CGGTAGATCA CCTGAGGTCA GGAGTTTGAG ACCAGCTGG CCAACATGGC
 7451 AAAACCCCTGT CTCTACTAAG CATAGAAAAA TTAATTGGGT GTGGCTGTGC
 7501 CTGCGTGTAG TCCCAGCTGC TTGGGAAGCT GAGGCCGGAG AATCTCTTGA
 7551 ACCCTGGAGG TGGGCATTGC AGTGAGCTGA GATTGCGCCA CTGCACTCCA
 7601 GCCTGGCGA CAGAGCAAGA CTCTGTCCTCA AAATAAAAGTA AAATAAAATG
 7651 AAGATAACAA CTGAAATTTC ACATTAACAAA TTTTTTGTA GCGACTGTGC
 7701 CTCCCTATGTT GTGCAGGCTG GTCTCAAACCT CCTGGCCTCA AGCGATCCTT
 7751 CCAAAGCACT GGGTGGGCCA CCATGTCCAG CCTGAAATTG TGCATTAAAA
 7801 AATTTCGCCG TTTTGGCTGG GCGAGGTGTC TCACGCTGT AATAGCAGTT
 7851 TGGGAGGCCG AGGCAGGAG ATCACTTGAG GTCAGTTCTA GACCGGCCTG
 7901 GCCAATGTGG TGAAACCTG CCTCTACTAA AAACACCAAA TTAGCTAGGC
 7951 GTGGTGGTGT GCGCTTGTAG TCCCAAGCTA CTGAGGAGGC TGAGACAAGA
 8001 GAATCGCTTG AATCTGGAA AAAGAGGTGG CCGTGAGGCCA AGATTGGCCA
 8051 CTGCACTCCA GCCTGGTGA CAGAGTGAGA TTCTGTCCTCA AAAAAAATAAA
 8101 AAATAAAAT TTCCCCCTTT AATCAAATTA AGTTAAATG AGGGATGTTA
 8151 GACAGTTTT AACCATCAA TATTTTAGTT TAGTTTTTTT TTTTAACGT
 8201 TGTCTTAAAG ATGGAAGTGC TTCAAAATCA AATCTCCCTT GCCAGTTCTC
 8251 TACTTGGCTT CTTTTTTTTT CTTTTGAGA TAGAGTCTCA CTTTGTCACT
 8301 GGAGTGCCTT GGCCTGATCT CGGCTCACTG CAACCTCCGC CTTCCAGGTT
 8351 TAAGTGATTC TTCCACCTCA GCCTCTCAAG TAGCTGGAG TACAGGTGTG
 8401 TGCCACCACA CCCGGCTAAT TTTTGTAGTT TTAGTAGAGA CAGGGTTCA
 8451 CTATGTTGGC CAGGCTGGCC TCAAAACTCCT GACCTCGTGA TCCACCCACC
 8501 TCAGCCAAAT TGCTGGGATT ACTTGTCGTGA GCCACGCGCC TGGCTTCTAC
 8551 TTGGCTTTA AAGGGAAATTG TGCTTTCTGA GTAATTCTAT TTCTCAGGTA
 8601 TCTTGGCTT TTAAATTCTG GAAGCAATCT TAATAATTAA TGTATGTGCC
 8651 CTGTAATCCC AGCACTTGG GAGGCCAGG TGGGCGAATC ACGAGGTCA
 8701 GAGATCGAGA CCATCCTGGC TAACACGGTG AAACCCCATC TACTAAAAAT
 8751 ACAAAAAAATT AGCTGGCGT GGTGGCAGGC GCCTGTAGTC CCAGCTACTT
 8801 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
 8851 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
 8901 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
 8951 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
 9001 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
 9051 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
 9101 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
 9151 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
 9201 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
 9251 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
 9301 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
 9351 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN
 9401 NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN NNNNNNNNNNN

FIGURE 3, page 3 of 21

9451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 9501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 9551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 9601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 9651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 9701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 9751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 9801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
 9851 NNNNNNNNNN NNCCAGGCTG GAGTGCAGTG GCACAATCTT GGCTTACTGC
 9901 AACCTCTGTC TCCCAGGTTC CAGCATTCTC TCTGCCTCAG CCTCCTGAGT
 9951 AACTGGGACT ACAGGGCGTCC ACCACCACGG CCAGCTAATT TTTATATTAG
 10001 TAGAGATGGG GTTTCACCAT GTTGGCCAGG CTGGTCTCCA ACTCCTGACC
 10051 TCAGGTGATC CGCCTGCCCT GGTCTCCCAA AGTGCTAGGA TTACAGGCCT
 10101 GAGCCACTAC GTTTGGCTGC TTATCAGCTT TTTACCACTT TGTCGCCACT
 10151 ACATTTGGA ATTTCTCTT GAGAATTAGG CAAAATGCC AGACTCCCC
 10201 CCGGCCCGC CTTTAGAGGG AGAGGGGAGC AATTAGACTA TTCCTTGTT
 10251 TCCCTATAGA AGGTGGGCT GAGATTACTG CTTTGATATC TGGAATGTAA
 10301 TTTAGGGAAG AAAATTAGG TCTTGGCCTT TCTTGGAAC CACCCCTGGGA
 10351 GTGTTGCGAGA TTATTAATAG GGTAATGGTG GAATGATATT CAGGGGAAAA
 10401 ATGGTCCTGA GGAGGCCAGAG AACTAAGTGT TAGTTGTTG GCTGACTGAA
 10451 ACAATGTGAGA GATAGGGTAC AGAAGAAGTA GGAAATAGTT TTCCCTGGTA
 10501 CTTCTGTGAC AGGTGGCTC AATTGGCTGG AACACCCCTAC ACTGCTTTAT
 10551 TAAATCCAAG GTTGTGATAG GTTCCAGTTA AGTTTACTGT GTTCTATGCT
 10601 TGTAGATTTC CTAATTAGGA CAAGTAGTGT TAAATATGCA TGCCTTTATT
 10651 CACAAGAGGG ACCATTCTT TGGAAACATC ACTTTTTAAT AATACTAGGT
 10701 GCTATTAGC ACTTACTCGG TGCCAGGCC AC TGGCTATGG TTTTTTTTTT
 10751 TTTTTTTTTT CGAGACATGA TCTAGCTCTG TCTCCCAGGC TGAGTGGTG
 10801 GTAGCACAGT CATGCTCAC TGCAGTCTCA ACCTCCTGTA CTCTAGTGAT
 10851 CCTCCTGTCT CAGCCTCTG AGTAACCTGGC ACCATGCCG GCTAATTNTT
 10901 TTTAAGAGAT GAGATGTCGC TATGTTGCCT ATGCTGGTCT CGAACACCTG
 10951 GGCTCAAGTG ATCCCTCCCG CCTGAGCCTC TCAAAGTGTG GGGATTACAG
 11001 GTGTGACCCA CCTCACTTGG CCATCTATGG TCTTTACATA GGGCATTNTG
 11051 TGCAGTCTGC ATCTCAAAC ACTGATCTTC AACAGTGAAA CTCAGTGAAT
 11101 TATGTAATTTC ATGTTTCCA AGAACAAATGA TGGAATTAA TTCTCTGAAT
 11151 GTATTTCTT TGTATAATAA TAGTACTTAA GTGGAATTAC TCTTGTCCCT
 11201 TTCTACTCTC CTTATAGATA TTTCTGGTA TCTTGATTTG GGACTGTTAC
 11251 ATTTAACCA TTTATGGTCG TGTAGCCATA CTCACGTTAC ATTTGATGCA
 11301 TCTGTCCTC TTGTGTCTAT ATACTCATAT AACATTTCG ATAAAGTTAT
 11351 AGGCACTCA CACCAAGGCT GTTCACTGAAC CTCAGATTAA GAATACTTGA
 11401 TTTAGGAGAT TGAAAACAGA AAAGAGAATG TTAACATATCA TTATCAATAT
 11451 TAAAATGTGA AAATCTGAGA GTGACAAAGC TTAGTTTAA ATCTGGTATC
 11501 CCAAACATCAT TTGAGTTTTT TTTTTTTTTT TTTTTTTTTT GAGACAAGGT
 11551 GTCGCTTTGT CCCCAAGGCT GGAGTGTAGT GGTGTGATCT TGGCTCACTG
 11601 CAACCTCCAC CTCCCAGGTT CAAGTGAITC TCCTGCCCTCA GCCTCTGAAG
 11651 TTGCTGGAT TACAGGCTGC GCCACCACGC CCAGCTAATT TTTGTATTT
 11701 ATAGTAAAAGA CGGAGTTCA CCTTATTGGC CAGGCTGGTC TCAAACCTCT
 11751 GATCTGTGA TCCCTCCGCC TCGGCCCTCCC AAAGTGTGG GATTACAGGT
 11801 GTGAGCCACT GTTCCCGGCC TAATTGAGT TTTAAAATGT GGAGTTAAG
 11851 ATGTTAGTCT TAAAGTGGGT TAGATGAAAT TTATAAAAAT AGTCAAATAG
 11901 CTAAATTAT AAAAGGCCAT TTGAAACAAAT TTGTGAAAT ATATAATGTG
 11951 GATAATTATG TAGTGTCTTA TGTGTAGATT GGTGGTTAGC ATCTGCCGA
 12001 TGAAGAGCAG TTGGATTCT TACTTACTAA AGCTAGTGAATCTGAACCTC
 12051 CAAATTAGGC ATCTTCACCA GGCTTTTTG AGCCGAGCTA ACTTACTCTC
 12101 TTTTTTATT TTATTTTTA ATTAAATTAAT TTTTTTTTTT TTTTTTTTT
 12151 TTGGTAGAG ACAGGATCTC CCCATGTTAC CCAGGCTTGT CTCTGGCTCC
 12201 TTGGCTCAAG CAGTCTCTT ACCTTAGGCC CCCAAAGTGC TAGGATTACA
 12251 GCTGTGAGCC ACTGCCAG GCTGAGCTTA TTCTCTACTA ACACAAGTGT
 12301 TCTAATTAA TTTAACAGT GATCACACT TTTCTTGTA TTTGGTCAGG
 12351 TTCTGGGTGC TAGTTTATAT ATGATTGAT TCATTCTGAT AGGGTTTTT
 12401 TGTTTTTTT TGTTTTGTT TTTTTGTTTT TTTTGAGACA GAGTCTAGCT
 12451 CTGTCGCCCA GGCTGGAGTG TGGTGGCTCG ATTCGGCTC ATTGCAACTT
 12501 CTGCCCTCCA CCCAGGCTGG AGTGCAGTGG CTCGATTTCG GGTCAATTGCA
 12551 ACCTCTGCCCT CCCAGGTTCA AGCGATTCTC CTGCCTCAGC CTCCTGAGTA

FIGURE 3, page 4 of 21

12601 GCTGGGATT A CAAGCACCCA CCACCATGCC CGGCTAATTT TGTGTATTT
 12651 TAGTAGAGAC T TGGGTTTCAC CATGTTGACC ACGCTGGTCT CGAACTCCTG
 12701 ACCTCAGGTG ATCTGCCTGC CTTGGCCTCC CAAAGTGCTG GGATTACAGG
 12751 TGTGAGCCAT CACACCAGGC CTCAAGAAC TTTTATTTT GAGACAGGGT
 12801 CTCACTCTGT CACCCAGGCT GGAGTACAGT GGTGAGATCA TGGCTTACTG
 12851 CAGCCTGGAC TTCCCAGGCT CTGGTGTACCC TCCCACATCA CCCCTGGAG
 12901 TAATTAGGAA TATAGACACA CACCCATGCC TGGCAGTTT TGTATTTT
 12951 TTCTTTTTTC TCTTTTTTG TAGAGACTGG GTTTCACATG TTGTATCAGG
 13001 CTGGTTTGA ACTCCTGAGC TCAAGCAATC CTCACTCTT GACCTCCAA
 13051 CGTGCTGGGA TTACAGGCAT GAGCCACTGT ACCTGGCCTT TTCTACATTA
 13101 AAAACTTTT ATTAAAAAAC CCAAATCTTC CTTGTGGTTG TATATAACATA
 13151 TATACATAGG TACACACATG GAGAATTTTA CCTTGGAGGA AGGCTTGGTA
 13201 AAGAAAATAG CCCTTGGGC CGGGTGCGGG GGCTGACGCC TGTAGTCCTA
 13251 GCACTTGGG AGGCTGAGGT GGGCGGATTG CCTGAGCTCA GGAGTTCAAG
 13301 ACCAGCCTGG GCAACACAGT GAAACCCCTGT CTCTACTAAA ATACAAAAAA
 13351 TCAGCTGGGT GTGGCAGCAT GTGCCTGTAG TCCCAGCTAC TTGGGAGCCT
 13401 GAGGCAGGAG AACTGCTTGA ACCCGGGAGG CAGAGGTTGC AGTGAGCCGA
 13451 GATTGTGCTA CTGCACTTCA GCCTGCGCGA CAGAGCAAAA CTCTGTCTCA
 13501 AAAAACAAA CAAACAAACA AAAAGGAAA ATAGCCTTTC TCTATCATCA
 13551 GAGTATATTA AGAGTTGAGT TTTTTTTCT GTTTTTAAA ATTTTGTG
 13601 TTTATTTAA ATTACAAAAC ATGGACTCTG CTTACAAATT AAGAAAATGA
 13651 CTCATGTCTA AACAAAGCATA ATCAATATAA CAGTTAATAC AAGTTAAATA
 13701 TTGTAATATG TTTACGGAAT AGCATGGCAA AATAGTGCAA AAGATTTGGG
 13751 GAAGGGGCCT ATAATTTCTG TTAACAGAAA GTTTAGTTA TGTGATTCA
 13801 ACTGGAGGAG AACAGAGCTC CCAGAAGGAC TCCAGAACAC TTGATGCTTG
 13851 TCTGAGTGGG GTCAGCAGCA CTGAGTTCCC ACCAGCCAGA AAGTTTGTGT
 13901 GTGTACATTA TTTCCCTTAA CTGCCACAAAT AATCCCCATGA AGAAAATGCC
 13951 CTAGTTTAC AAACAAGGAA ACAGAGGCAG AGAAGAGTTA AATGACTTGC
 14001 CCAAGGGCAT TCAAAGTAAG CAACTGAATT GGAATTTTAA CTCAAAGGCT
 14051 TGGATGTCCC ACTACAACAA ATAGGCTGTT TCTGCTTAC TACATGTGCT
 14101 TACTTCTAAG AATTAAACAT TTTAGGCTGG TTGTGGTGGC TCACCTCTGT
 14151 AATCTCAGCA CTTTCGGAGG CTGAGGTGGG TAAATCACTT GAGCTCAGGA
 14201 GTTTGAGACC AACCTGGCA ACATGGTAAA ACCTCATCTC TACCAAAAAAA
 14251 AAAAAGAAAAA CTAGCTGGAC GTGGTGGCAC GCGCCTGTGG TCCCAGCTAC
 14301 TCAGGAGGCT GAAGTAGGAG GATCGTTTGA GCCTGGGAGG TGGAGGTTGC
 14351 AGTGAGCCCA CATTGCATCA CTGCACTCTA GCCTAGGTGA CAGAGTGAGA
 14401 GCCTATCTCA CACACAAAAA AAAGAATTAA AAATTTAGT CAAGTAATTA
 14451 GGCACTAACAA TTTTGTGGTC AGTTACTTTA CGAATTCTG GTGGAGGCC
 14501 TGATGTGGTG GCTCATGCCT GTAATCCCAG CACTTTGGGA GGCTGAGGCA
 14551 GGAGGATTGC TTAAGGCCAA GAGTTCAAAT CAGCCTGAGC AACCTAGTAA
 14601 GATCCCCCTT CTGCAAAAAA TTTAAAAAATT AGCTGGCCT GGTAGTGTGC
 14651 ACCTGTAGTC CCAACCCTT GGGAGGCTGA GGTGGGAGGA TTGCCTGAGG
 14701 CCAGGAGTTT GAGACCTGGG CAGCATATGA AGACCCGTCT TCTAAAAAAC
 14751 TAAAAATAAA AAATAGCCAG GTGTGGTTGG TGTGCTTGTG GTCCCAGCTA
 14801 CTCAAGAGGC TGAGGCAAGA GGGTTGCTTG AGCCCAAGAAG TTGGAGGCTG
 14851 CCGTGAACGT TGATTGCACC ACTGCACTTC AGCCTGGGTG ACATAGCAAG
 14901 ACCCTGTCTC TGTGGTGGTG GTGGGTGGGG GTGGGGGAAG GGATTTAAGA
 14951 AGGGTTTGTG AGGTATGTAT TATTTATAAA TGGGTTTTA ACTTTACCC
 15001 TCACATCTTG GTTCAAATT AATTGTATCC ATTCTCAGTT TTTCTGTCTT
 15051 GCTATATATT TAAACTTGGG GACTTAGAGG TCATGGATGT CTTTCTATGA
 15101 AAAGCAAATG AAGCAGAGGG CTGCCCTCTC TTGCTGTAGA GGGCACACTT
 15151 GCTGCAGAGC ATGTTACTGT TTTATGCATT GCTAGGCTTT GGGAGTTGTG
 15201 ACTTGTATGA TCATAGTACT TACAACATTAGTGGTGGCAAT TTTTAAACTT
 15251 TAACTTTAGA TTATATATGT AAACCTCTG GTTCTTTGT CACTGATAAT
 15301 CTGAACAGAA GCCTTGGATA AATAATTGG AAGTTTTGT CTGAACCTCT
 15351 GAAATTGTGA TTGTATCTC ATGGTTTGG TGGGAGGAAG GAGAAATAAC
 15401 AATGGCCACT TACTGTGCTT CTGTATGTGC CAGACAGTAT GTGCTAGATG
 15451 TTTCAGAACAC GTGATTGTGA ATCCTGACAA GAAGCCTAAT TGGGTGGTAG
 15501 TGGGTGCTAA TTGAACCTTA TAGATGAGGA AATTGAGGCT CATGGTGGTA
 15551 AGTGAATAAC TTGCACCAAG ATCCATGGC TGGTATGCAG TAGAGCCTCA
 15601 ATTCAAGTAC GGGCTTCCA GGTCCAAACCC CATGCAGGCT TTGAGAGGTA
 15651 AGGAGGTAGA GAACGTTGAC ACCCCCTTCT TGGTGTGTTT TTCAGCAAAT
 15701 ACTTGTATGC ATATTAAGA CTGTCTACCC TTTTGTCTAC TGTGTCACT

FIGURE 3, page 5 of 21

15751 TGCTGCTTCC TTTGGTACTA CCCAAATTTC TTTCAGCATT TCAGCTTG
15801 ATTTTTATTT TTATTTTATT TAATTTATTT ATTTTTTGA GATGGAGTCT
15851 CACTCTGTTG TCCAGGCTGG AGTGCAGTGG CGTGATATCA GCTCACTGCA
15901 ACCTCTGCCT CACAGGTTCA AGCAATTCTT CCTGCCCTAG
15951 AGCTGGGACT GGAGGGCCCC ACCACCACGC CCAACTAATT TTTGTATTT
16001 TAGTAGAGAT AGGGTTTAC CTTGTTGGCC AGGCTGGTT TGAACCTTG
16051 GCCTCAAGTG ATCCACCCAC CTCGGCCTCC CAAAATGCTG GGATTACAGG
16101 CATGAGCCAC TGCACCTGGC CAGCTTGAA TTTTTAGAAT ACTGTTCTAA
16151 ACAGAACTAT ATTGGAACCT GGAAAATTAA TCTATTGTCT CTAAATACCA
16201 AAGAAAACA TGTAATTTA GTGGTTGATT ATGGGAACAA TTTTTTTAA
16251 GATGGTTCAT CTGAATGGGA AGCATTTTT TTTTAATTGC TTGACTATTT
16301 CTTTAAATTT GGAGAAAAGA CCATTGCCCT CTCAGATTTC TGGTAATTGG
16351 TCACATTGAT CATTATATT GACTGACAGG CTGCTTGTC CACAGCTGAA
16401 GGATTGTTA ATTTTTTTA AATTATAAGA GTAATATGTG CTCACTGTAA
16451 AATTACAGT ACAGAACAT ATGAACTAAC TAAAAGTTCT TACCTCTTGT
16501 CTCCAGCAAG GAGTAAGTGT TTCAACCTGA AGGTTGGTT TGAATTGTGT
16551 TCTGTGGAGC GTACTTAAAG TGAGTGAAGA AGAAAAATTAT ATGTCATCA
16601 TGATCATGTC AGCTGAAGTT TTTATTGTTT CACCCCTAA AGTTATTAA
16651 AATAGTATGT AGTTAGTAG TCTTGATAAT TTTCCCTAA GATTATTGG
16701 CCAGTATATC AGGATTTGT TTTAAATTG ATATGTGAGC TTAGTTTTAT
16751 GCTATTTCA AATAAGACAT TTAGAAGAAG ATAAAATAAC ATTCCGTCT
16801 TAGTCTGTTT TCTGTGCTA TAACAGAATA GCACAGACTG GGTAATTAT
16851 AAACAGTAGA AGTTTATTG GCCTGTGGTT CTGGAGGCTG GGAACCTCAA
16901 GAGCATGGTT CTGCCCTTG TGCTGTGTTA TCATATGGTG GAAGGTGGAA
16951 AGGCAAGTGG GTATGTCAG ACAGAGAGCA AGAAGGGGCT TGAACACT
17001 TTTATAACAG AGTGAETCCA GAGATAGCTA ACCCACCTTT GAGAGAATGC
17051 ATTAATCCAT TCATGAGGGC AGAGCCCTTG TGACCTAACACCTCATT
17101 AGGCTCTGCA TCCTTAAACT GGTTTTTTTG TGTTTTTTTG TTTGAGACG
17151 GAGTCTCGCT CTGTTGCCCA GGCCGGACTG CGGACTGCAG TGGCGCAATC
17201 TCGGCTCACT GCAAGCTCCG CCTCCCGGGT TCACGCCATT CTCCGCCCTC
17251 AGCCTCCCGA GTAGCTGGGA CTACAGGCGC CCGCCACCGT GCCCGGCTAA
17301 TTTTTGTAT TTTTTTAGTA GAGACGGGGT TTCACCTTGT TAGCCAGGAT
17351 GGTCTCGATC TCCTGACCTC ATGATCCACC CGCCTCGGCC TCCCAAAGTG
17401 CTGGGATTAC AGGCGTGAGC CACCGCGCCC GGCCCCCTT AACTGTTGT
17451 ATTGGGGATT AAGTATCTAA CACAGGAAC TTGGAGGATA CATTAAACC
17501 ATAAGAATTTC CTGTCATGCA ATGAATCCA TTCTAGATGA AAGAGAATGA
17551 ATTTAGTTTC CATTGAACCTT TATAAATAGG CCTTTCTAA GGTAATTACA
17601 GCTGATATTA TAAAATTAT ATTGTTTTT ATAAATTGT ATTGTATTT
17651 CTGTTGTAC AAATACAATT ATACACTATA GTTCTCTGCT GTTAGATTT
17701 TTTTCTCCT TAGCATGTTT CCAAAGGGTG GAATGTGAA AGTGGGTTA
17751 ATGTCATCA GCTTCTTTT GTAAAGTGT CATTGACATG TGAACCTTGT
17801 CTGAGAATCT AAATTATTT TCATGAAAGA AGAAAACAGT ATATTCTCAT
17851 TTAACCCAGA ATTTAACCTC ATATACTTGT GGCTGTATTG GGAGTATGCC
17901 ATTGCTGTCT GTTTACAACC TGACCTACTC TACCTACTTA GAAGTAATT
17951 GTGTTATGAT AGGTGTGCTG TGCTGACATA TGCTGAACAT ATTTGTAAGG
18001 GTGTTAACTC ATTGAATAAA ACGCTTTCT CCTCCTTCA AATAACATT
18051 TTTATTTCTG GTTATAAAAG TCATACAAGC TTACTGCAGG TTGTTAAAAA
18101 GGTATAAAAGA AGAAACCGTC ATCCATTAT AATCCCTACAG TTTAGACTTC
18151 CTGCTCCAGC CTCTCAGAGT GCTGAGATGA GCTAGCCATG CCCAGCCCT
18201 CAAAAGATT TTTAAAAAAC AAAAATGAGG TTATACCTTA AAAAATTCTA
18251 TATTCCCTTC ACATAACAGT GTTATTTTG AGGTTTGTAG ATTTCCAGTA
18301 GCATTTAGA TTCAGAAACA AGCTGATTCA TCCTCTACTT TCTACTTTAG
18351 GCAAGAAAAG AATTTCACCT AAATAGAATT TTGAACGTGAA AATCTGTTT
18401 TCTAACCTTT TATTTAAAGA ATATTGTTCC ATGCTTTCAC AGTACTGACT
18451 TTTAATTTTT ATATTTTTA TTTTATTTAT TTAGAGATGG GGGCTCACT
18501 CTTGTTGCCT AGGCTAGAGT GAGTGCATG TTGCTATTCC TAGCTCACTG
18551 CAACCTGAA CTCCGGGCT CAAGTTACCC TCCTGCCCTCA GCCTTCTAAG
18601 TAGCTGGGAC TACAGGTGTG CACCACTGCA CCAGGCTTT TTTAAAGGCA
18651 TAGAAAATGG TAGTGCTTGC ATACAAAAAT GGCCTAGGTA CACACATCAG
18701 CGGACATCAA GACTATGTTC AGATCATAA TGTACATATA TGTACCGATG
18751 CCATTTTGC ACGAAACAA ATAATGGAAA TTGAACCTCA AACTGAAATT
18801 TGAAACAAGG GTTCTGGGGT GGGCCCTCTT GCTGATTGT AATTGAATGT
18851 ATAGTTCAAT TTTCCCCAT CTGTTAAGCA AAAGACAATT CTAATGTTAG

FIGURE 3, page 6 of 21

18901 CAAAAATCCA CATATCCTGT CATTGATCAT TTTTCCTTA ATTTTCTTA
 18951 AGAGATGGGG CTTCTCTCA TGTTGCCAG GCTGGCTGG AACTCTGG
 19001 CTCAAATGAT CCTCAGCCT CAGCCTCCC AAGTGCAGG ATTAATAGGC
 19051 ACAAGCTGCT GTGCCTGGCC CTGTCATCAG TCATTTAATC TCATGCAAAC
 19101 TGAGTAGAAT AAAACTCGTC CTTACTGTAC CTTATTGCTT TTGTTTATT
 19151 GTTGGAACCT CCAATATTGC GAAAGTAGAC CAAAAGTTGA CTTATAGGAA
 19201 AAAACTGATAG CAAAAATAAT TTTCTCTTG TTGCTGTATT TCATGCCAC
 19251 CATCCAGTTG TAAAGCTTA CTGTTAATT CTCTCAGCCT CCTCCTTTCT
 19301 GTCCAGGCTT ATTCTATGCC ATTCTTACCT TAACTGTTT TAGCTTCTC
 19351 ATAGAGTGA CTTTTAAAT TAAAATAAA TATCTGCTCG TAGTATTATA
 19401 AAATTCAAGC AGTTCAACAG AATTTTCAC TAATAGAAAT ACTTGTACCT
 19451 CAAAAGCAGC TTTATTTAC AAACCCAGCC CAATTGTGA TTAGATTTAA
 19501 CTTGAGAAAA CATGAAATGT CTCTCATATT GTTTAAAAAT ATCATAAGTG
 19551 GCTGGGCACG GTGGCTTATG CCTATAATCC CAACACTTG GGAGGCTGAG
 19601 GCAGGTGGAT CACTTGAGGT CAGGAGTTG AGACCAGCCA GGNNNNNNNN
 19651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNN
 19701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNN
 19751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNN
 19801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNN
 19851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNN
 19901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNN
 19951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNTTC ACCATGTTGG CCAGGCTGGT
 20001 CTCAAACTCC TGACCTCAGG TGATCCACCT GCCTGGCCT CCCAAAGTGC
 20051 TGGGATTATA GGCTGAGCC TCGCCTGGCC TCCTCATATA TTTTAACCT
 20101 TTATAAAAAC CTTTCTAAA ACCCTTTTA TTTGAACTA AATTAGATT
 20151 TACTGAAATT GTGAAATCAA TGTGGAGTT TCCTATACCC TTCTTCCGC
 20201 TTTCTCTAAT AGTAAACATCT TACATACATG GTACATTGT CCAAATTAAG
 20251 AAATAAACAT TGGTACAGTG TTAACTATAG ACTTAATCTG GTTCTCTAA
 20301 TTTTTCACT AATGTTCTT TTCTGTTCTA GGATCTAATT CAGTATACCA
 20351 TATTGTTATT AGTTGTAGGC CATGTTAGCC ACCTTCAATC TGTGACAGTT
 20401 TCTCAGTCTT TCCTCTTTT TCGTTATCTT GACAAGTTG AAGAGTGTG
 20451 ATAGGTATT TATAGAATGT CCGTCAGTT TCTGTCAGTT TGTATTTGTC
 20501 TGATGTTATT TTTTTTTTTT TTTTGAGATG GTGTCCTGCT CTGTCGCCTA
 20551 GGCTGGAGTG CAATGGCATG ATCTTGGCTC AATGCAGCCT CCACCTCCGG
 20601 GGTTCAAGTG ACTGTCCTGC CTCAGTCTCC CAAGTAACTG AAAACTACAGG
 20651 CATGTGCCAC CACGCCCTGGC TAATTTTTTG TATTTTAGTA GAGAAGCAGT
 20701 TTCACCGTGT TGCCCAAGGCT GGTCTCGTGC TCCTGAGCTC AGGAATCCA
 20751 CCCGCATTGG CCTCCCAAAG CGCTAGGATT ACAGGTTGTGA GCCACCATGC
 20801 CTGGCCAATA TTTTGAGGGA TATACTTTGG TGAGGTCATG CAGATATCCT
 20851 GTTTCTCCTT AGTTTATCG ATTAATTAGT CATTATCCA GTAAATCTTC
 20901 CTTGCAGCAA TTATTTTTT TTTTCTTTT TTCTCTTAATT TTTTTTTAA
 20951 GAGATGGGAT CTCACTCTGT TGCCCAAGTT GGAATGCAGT AGTGAGTTCA
 21001 TAGCTCACTG CAGCCTAACTCCTGGGCT CAACTGATCC TTCTGCCTCA
 21051 GCCTCTCAAG TAGCTGGGAC TACAGGCATA GACCACACCA CCCAGCTAAT
 21101 TAAAAAAAT ATTTTTAGAG ATGGGGTTTG TGCTATGTTG CTCAGGCTGG
 21151 TCTTGAACCT GCTGCCCTCA TGTGATCCTT CTACCTCAGC CTTACAAGTA
 21201 GGTGGGAATT ACAGGTGTGA GCCACCACAC CCAGCATTGC AGCAATTATT
 21251 AATGTAGTGC TACTGGTCAT TTCTGTTT TCTCATTCT TCAGCATGTG
 21301 TTATTGACTT GTCTCTTCCC TCCCATTAT AATCATTAT ACTGCTATGA
 21351 ATTCACTGAGT ATTATTGTTG TGAGTTATAA TCTAATACGT ACTTAAATT
 21401 TTTTGTGCTT CAAATTGTTG TGGCTTGGCC ATTTTTTTT TTTTTTTTG
 21451 AGACGGTCTC GCTCTGCTGC CCAGGCTGGA GTGCAGTAGC GCCATCTCTT
 21501 CTCACTGCAA CCTCCACCTC CCGGGTTCAA GCGATTCTCC TGCCTCAGCC
 21551 TCCTGAGTAG CTGGGACTAC AGGCCTGTGC CGCCACACCC GTCTAATT
 21601 TTGTATTATT AGTAGAGACA GGTTTACCC ATGTTAGCCA GGATGGTCTC
 21651 GATCTCCTGA CCTCGTGCAT TGGCCCCCTC AGCCTCCAAA AGTGCCTGGG
 21701 TTACAGGTGT GAGCCACCAA GCCCGACCGG CTCCGTATC CTTTAACAT
 21751 GAGGTGCTGT CATCATTTT TCCCCCTAAT ATTTGGCCA AAAATGTTAA
 21801 TCAAGGATGG CACAAATTTT CTGTAGCTGT ATCTCACAAAT GAAAGAGGCC
 21851 TGATTAAGAA TGAAAAGTAA AATGTTCTC TGATCTCTTA GCACATGCTT
 21901 TGTAAAAGGC ACAGTGCAG ATCCTTGTAT ACGTAGATGA GTAAGTCAGC
 21951 TTACCTTCCA CACCCACAGA TAGCTATGTC AAACGTAAGG GTGGAGAAC
 22001 ACAGACCCCA AACTCTCGA GGGTAGAAAA TATGAGGTTA TAGTAGATTA

FIGURE 3, page 7 of 21

22051 GAACTACAAA AAGCTAGAGG AAGTTCTGAA CTGGAAACAG TGGATAGGAT
22101 TTACTAGAAT AATTACGAG GGTGACAATT GTAAATCTTC ATAGGTTCT
22151 TTTTTTTCCT TTCTCTTTT TTTTTTTGA GATGGAGTCT CGCTCTGTTG
22201 CCCAGGCTGG AGTGCAATGG CGCAGTCTCT CCTCACTGCA ACCTCCGCCT
22251 CCTGGGTCCA GGTGATTCTC CTGCCTTAGC CACCCAAGTA GCTGGGATTA
22301 CAGGCATCTG CCACCATGCT GAGCTAATT TTGTATTTT TTTTTAGTA
22351 GAGACGGGGT TTCACCATGT TGGTCAGGCT GGTCTTGAAAC TCCTGACCTC
22401 AGGTAATCCA CCCACCTTGG CCTCCCAAAG TGCTGGGATT ACAGGTGTGA
22451 GCCACCAGCGC CCAGGCCAAAT TTTTATTGGT TTCTAAACTA GCGTAATTIA
22501 GTTTTTTCA CTTAAGTCAA AATTATATTA TTGTAGGATA AAAACTTAGT
22551 GATCCAAATT CATGAGGAAT GAAGAATAAA TACATTAAA GTCTTACCAT
22601 TTGCTAAATT AGTCTTGGCT CTTTGTACCA AAATTCTGTC CTTGTGCTCT
22651 GTAATTATAT ATTGTATAT TTTCTATCAA CATTTTACT GTGTGGTGT
22701 TTGTAAATTA TAAAAACGTT TAAAGCAAA CTCAGAACAA TGAATTCTCA
22751 CGAATATTCA GTATATTAC AGTTGAGAAA TAAACTACTT CTGTAGTAGG
22801 TAATTTAAAAA TGTCCCAATG CAAGTTAACG TGTCACTGAT CACGCTATT
22851 AGGTGTGTGT CTTTGATAAG GGGAGGTGGG GAAGTTGTG GGTTTGATT
22901 TATTTGCCCT TCTCATGTGA CTGTGTGAT GTTAGTAAAC AAATGGTTG
22951 CGAGAGAACC AGTAGTCTT TGCAAAGATT GTCTTATACA GAGCACTCAA
23001 TTCTTCATAT TATTATAAT GGCTTTAATT TAAGCCTTAA ATTATTAGAA
23051 ACTCATAAAT AATTTTTTA TTTGTTTTTG TGAGATGGAG TTTCGCCCTT
23101 ATTGTCAGG CTGAAGTACA ATGATGTGAT CTTGACTCAC TGCAACCTCC
23151 GCCTCTCGGG TTCAAGTGTAT TCTCCTGCCT TTGCTCTCCA AGTAGCTGGG
23201 ATTACAGGCA TGCGCTACCA TGCGCTGGCTA ATTTTGATT TTTAGTAAAG
23251 ACAGGATTGC ACCATGTTGG CCAGGCTGGT CTCGAACCTCC CAAACCTCAGG
23301 TGATCCACCT GCTTCGGCCT CCCAGAGTGC TGGGATTACA GGCTCACTGA
23351 GCCACTGTGC CCAGCATAA TGCGTTAAA TAAGAGTGT ATATTGTAA
23401 AACTTAAAAA AATGTAGTGG TTGAAAAAAGG TAATTTAAA AGAATTGACT
23451 ATTAATTCTC TGAAACCATA ATGTAACCTG TAGTGCAATT AGGAAACCTT
23501 CATGTTTCTT TCTTCTTTT TTTTTTTTTT TTTTGAGAT GGAGTTTGC
23551 TCTTGTGTC TAGGCTGGAG TGTGTGATGT CAGCGCACTG CAAACCTCTGC
23601 CTCCCTGGTT CAAGCAATT TCCTGCCTCA GCCTCCCGAG TAGCTGGGAT
23651 TACAGGCAGC TGCCACCACA CCCAGCTAAAT TTTTGATT TTAGTAGAGG
23701 CGGGGTTTCA TCGTGTGTCCTG CTGGCTGGTC TCGAACTCCT GACCTCAGGT
23751 GATCCACTGC ACCTGGCCCC CGTTCATGTG TTTTAAAGCT TTATGGTTGC
23801 TCTGAAATAG AGTTGTTGAT TTTTTTTTTT TTTTFGAGAC TCCTCTTTG
23851 CCCGTGCTGG AGTGCAGTGG TGTGATCTGA GCTCACTGCA ACCTCCACCT
23901 CCTGAGTTCAGCAACATTCTC ATGGGTCAAGCTCCTCAAGTA GCTGAGAGTT
23951 AAGCTGCCCA CCACCATGCC TAGCTAATT TAGTATTAGT TTAGAGAGATG
24001 GGGTTTCACC GTATTGGCCA GGGTGGCTCTG GAACCTCTGA CCTCAGGCAT
24051 GAGCCACTAC GCCTAGCCTG GGTTGTTGAT CTTTAAGGTG ATACCTCAGG
24101 CAACATCTGA GGCCCAGTAC AGTCCTTTAC TTCAACTGGC TCCAGTACAG
24151 CAAATTCAAGG GAATGTTTT GAGTGTGTTAC TGGATGCCTG GCGTGGAGTT
24201 CAGGGAGATT GGTACATTGA GTCCAGTTGT TGTGTGAAA CTTCTGTTTA
24251 AAAACCTCCC TACTAAGTCC CAGCTACTCA GGAGGCTGAG GCCTGAGAAT
24301 CACTTGAAACA CCTGGAGGCA GAGGTTGCAAG TGAATCGAGA TCGAGCCACT
24351 GCACTCCAGC CTGGCGACA GAGTGAGACT GTCTAACAAAC AAAAACAAACA
24401 CCCCCCAAAA AACCAACCTA CTATGGTAGT ATCAATGCTG TGATAGTCTT
24451 CCTTTCTTCAGCAACATTCA TACAGGTAAA TTCTTAAACAT ATACTCATTTG TTAATGTTCA
24501 GTGTTCAAGTGG TCTTAAAGAG TATTGTTGGG CAGGCACGGT GGCTCATGCC
24551 TGTAATTCTCA GCACTTTGGG AGGCTGAGGT GAGCAGATT CCTGAGGT
24601 GGAGCTTGAG AACAGCCTCC AACATGATGA AACTCCCGTC TTTACTAGAA
24651 ATACAAAAAT TAGCTGGGT TGTTAGCACA TGTCTGTAAT CCCAGCTACT
24701 TCAGAGGCTG AGGCAGGAGA ATTGCTTGAA CCTGGGAGGT GGAGGCTGCA
24751 GTGACCTGAG ATTGCTTCAC TGCACCTCCAG CCTGGGCAAC AGAGCGAGAC
24801 TCTTGTCTCA AAACAAACAA ACAAACAAAG AATATTGGG GCCAGGCATG
24851 GTGGCTCACA CCTGTAGTCC CAGCACTTTG GGAGGCAAG GTGGGTGGAT
24901 CACTTGAGAT CAGGAGTTGG AGACCAGCCC GACCAACATG GCTAAATCCC
24951 GTCTCTACTA AAAGTACAAA AATTAGCTTG AGCAACAGAG CAAGACTCTG
25001 TCTCAAAAAA AGAAAAGAAGA ATATTGTTT TAATTAAGAA GGAACCTTAT
25051 CAATAGTAGT AAAGTCAGCC AGCTGAACCTG CCAAGTACAA ATTGTTGGTA
25101 TTAGGTATCA ATCATTATT AAGGATAATA TTCTACAATA GCGATCTTTT
25151 TAAAAATTAA AAAATCTCAA ACTGGAAAGG ATGTCTAGTT CATTCTATGC

FIGURE 3, page 8 of 21

25201 TTCAAGTCCCC TCTTCTGATT TACTTGTAA GAAGATTTT GTTCCTTCT
 25251 CTGACTTCTA TTTTGCTGCT GACTGGCACT TGGGATTTT AAAAAATTAT
 25301 TTTCCTCATA TATAATTAAA GACAATAAGT ATAACAATAA GTATAATATG
 25351 GTAATTGCT AAAACCCAAA CAATGTTTA AGTAATGCAT ATCATTATGT
 25401 AAACCTACGT AATAGTTGAA TATTCACAAA GATAATCGCT TATAGAAGTT
 25451 TTATATCCTC TCTTCTTGG CAGTGCATT AAAACAAAAA AAATAAGTTT
 25501 TATGTCITGT TTACATGTAA ATAATTAA TCTAAATTGT GACGTGGTTT
 25551 TCACTTTAGC ATATTGTTGA AAGTAAATCA AAAAGGACAA AATACAAAAT
 25601 CATGTATATC TTCTACAAA ACGATATATA AATTCTAAGG TTTTGTCTC
 25651 TTTGAAATTG CTTAAAAGAA TGCATAGAAC TGGTGTCTGA GTGGGGAGG
 25701 ATCTATGAGG GATTCCCTG GAGACCGTGG GTGAATAATA ATGTTGTCTT
 25751 AGTTCCATGA AGGAATCTCT GGGGATAGTT TTTGAGTTAG GCCTGGCAAT
 25801 GTTAGAGATA CATAAAGAGA GCCTGTTTT ATCACTGGGT GCGGTGGCTC
 25851 ACACCTGTAA TTCCAGCACT TTGGGAGGCT GAGGCAGGCA GATCATGAGG
 25901 TCAGGAGATC GAGACCATCC TGGCCAACAC GGTGAAACCC GTGTCTACTA
 25951 AAAATACAAA AATTAGCTGG GCGTGGTGGC GCATGCCTAT AATCCCAGCT
 26001 ACTCGGGAGG CTGAGGCAGG AGAATCACTT GAACCAGGGA GTTGGAGGTT
 26051 GCAGTGAGCC GAGATCGCAG CACTGCACTC CAGCCTGGGT GACAGAGCAA
 26101 GACTCCGTCT CAAAAAAA AAGCTTGGTT TTCAATGGTT CTGAAAATG
 26151 CTTTAATACA AGTGTAGAGT GTTAGTCAGG TTTTGCACTT GGATAAACAG
 26201 CCTGTGAATT TATCACATT CTAGTTTATA ATATGGCTT TCAGAAGTTA
 26251 TATGAACATT GTTTTGACGG GAGAATTCAA GCTGGATGCT AGAGAAGGAT
 26301 CGTGAGAAC CCTTCATTGG AGGAGTGTAA TGAAATTATT TGATCTTGG
 26351 ATTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTGAGAC AGAGTTTCGT
 26401 TCTTATTGCC CAGGCTGGAG CTGGAATGCA GTGGCACGAT CTCGGCTCAC
 26451 TGCAACCTCT GCCTCTGGG TTCAAGCAAT TCTTCTGCCT CAGCCTACCA
 26501 GGTAGCTGGG ATTACAGGCA TGCGCAACCA TGCCCCAGCTA ATTTTTGTAT
 26551 TTTTAATGGA GACGGGGTTT CACCATGTTG GTCAGGCTGG TCTTGAACTC
 26601 CTGACCTCAA GTGAATGCC TGCCTCAGCC TCCCAAAGTG TTGGGATTAC
 26651 AGGTGTGAGC CACTGCGCCT GGCCTGATCT TAGAATTGAA AGGAGAGACT
 26701 AATATTCAT GGGCAAAAC AATGAAAAGT TACCTTCTG TATTCTAATA
 26751 CTATAGAGGA GTGGGATTTA TTTAGAATGT TTTAAGTATC TTGGGCAGTC
 26801 CAAGAGTGC G TATCACTTAT TTTTCTTTTC CTTCTTTCTT TTTAAGTGG
 26851 AGTTCACTGA TGTAGAGAT CATAGGTGGC ATTGCCTACT TTTTACATAA
 26901 TTTTATCATG TTTAGTGTAC TGTCAGAAGG GCTGTGGCTG TTTGCAGTTT
 26951 TGGCTTAAGC CATGCATGGG CTTTATAGGA GATGTAGTCT TCACAGTGAG
 27001 TTGTTATTG TAGCTGTGTT TTTGTTTTTG TATAGCTTAT AGCAATGCAG
 27051 TGTGCTTTT ATTAACATCA TTTTCTTTTT CTTTTGCAG TGATTATTAA
 27101 TTCAAGTTAC TTCTGATTGG CGACTCAGGG GTTGGAAAGT CTTGCCTTCT
 27151 TCTTAGGTT GCAGTAAGTT GAAATTGAAA TGTCTTACA ATTAATGGTA
 27201 CAATTAATGC TATGTATGTT TTCTAGGTAG ATAAAAATTAA ACAGTTTTAT
 27251 TCAGAATAAG TTAATTCTTC CAGAATTAT ATATTTAAAG ACTCCAAATA
 27301 TACATCCCCA GTGGTATCTT GGACTGTTAA ATAGAAAAAT ATTGTTGCTC
 27351 TTAAAAGAAA TTCAGTGAAG TCTGGTTATA AAGTCAGAAT GTCTAATACT
 27401 TTGGTCAGA GTCAAACAGC AGTTCCAATA TAGGCAGCAA GTTAAAGGGG
 27451 TAGTTGGTGG CCTGTGTGA AAGCGACTTG ATGAAAATAA ATCTTTAAAT
 27501 TAAACTTAG TAGAATAAAA AGAAAAAGCA GAGCCAGGTG ACGCAGTGG
 27551 TCATGCCCTGC AGTCTCAGCT ACTCAGGGTG CTGAGGGTGG AAGGATCACT
 27601 TGAGTCTAGG AGTTTGAGA CCAACCTGGA CAACATAGCA TGACTCTGTC
 27651 TCTGAAAAAA AAATTTAATA AAAGAAAAAG TAGGGTCTTG GACAAACTTC
 27701 GTTGGCCAAT GGCATAGTTC TAAATGCTGA AGCTGACAGA TAAAGGACTT
 27751 TTGACTTIAAC AGAATCCACA GTGCTCTCA TAGTCTTTAT CAACTACCTT
 27801 TAAATTTAGC ATGTTCTTG GCGAGTGC GTGGCTCACG CCTGTAATCC
 27851 CAGCACTTGG GGAGGCCGAG AGGGGGGAT CACAAGGTCA AGAGATTGAG
 27901 ACCATCCTGG CTAACACGGT GAAACCGT CTCTACTAAA AATACAAAAA
 27951 ATCAGCTGGG TGTGGTGC CACGCCGTGA GTCCCAGCTA CTGGGGAGGC
 28001 TGAGGCAGGA GAATCGCTG AACCCAGGAG CGGGAGGGTG CAGTGAGCTG
 28051 AGATGGTGC ACTGCACTCC AGCCTGGCAA CAGAGCAAGA CTGCTCAAA
 28101 AAAAAAGAAA AAAAAATAAA AAAACAAATT AGCATGTTTC CCTTCTAGAG
 28151 ATCATTGTTT CTCAGAGCAT GGACCAAAGA CTCCTGGGG TTACCAAGAC
 28201 CCTCTCAGGT AGCCCATGAG GTCAAAATAT CCTAATAATA CTAAGATGTT
 28251 AGTATTTGTA AGGAAATATT TACTTGGTAA TAATACTAAT ATAAAAGATG
 28301 TTTGCCTTT TCAGTGATGA CATTGGCTCT GGTACAAAAG CATGTGGGTA

FIGURE 3, page 9 of 21

28351 AAATTGCTGC TGGCTTGGTA CACATCAAGG CAGCGCTAAG CTCCAAATTG
28401 TACTCATGGT GATGGCATTG TTTACCTCTG TGCCCTCACA GGAAACAAAAA
28451 CAAGCCGTGC CATTTTTATT GAAGATTGTC CTTGACAAAAA CAGTTAAAAT
28501 GATTAATTTC TGAAAAATGT TGATCCATGA GTATTCCCTT AAAAATATTT
28551 GTGAAGAAAT GGGAGTTC AATAAAACAA TGTTTTTTT TTGTTTTTT
28601 TTTTTTTTTT TTTTGAGACA GATTCTGGCT GTGTTGCCAA GGCTAGAGTG
28651 CAGTGGCGTC TGGCTCCCAG GCTCAAGCTG TTCTCCCACT TCAGCCTCCC
28701 AAGTGGCTGG GACCTCCAA GTGGATGCGC CATCATGCCT GGCTGATT
28751 TGTATTTTTT TGTAGTACA AGGTCTCACT GTGTTGCACA GGCTGGCTC
28801 AAACCTCTGA GCTCAAGCGA TGCATGTGCC TCAGCCTCCC AAAGTGTGG
28851 AGAAAGCACT TTTTACTGCA TACTGGCTA TGTGTTGGTT ATTTGGAGA
28901 AAAGAAAAGC ATTTGTAGTT TTTGAGTT TAAGCTGAGC TAACTGCTTT
28951 ATTTTTTCT GTGGAACACC ATTCTTTT TTTTTTTGAA GATGGAATAT
29001 TGCTTTGTT CCCAGGCTGG AGTGCAGTGC CACAATCTCG GCTCACTGCA
29051 ACCTCCGCTT CTCGGGTTCA AGCAATTCTT CTGCCGTAGC CTCCCAAGTA
29101 GCTGGGATTA TAGGCACCTG CCACCAAGCC CAGCTAGTTT TTGTTTTT
29151 AGTAGAGATG GGGTTTCACC ATGTTGGCCA GGCTGGCTC GAACTCCTGA
29201 CTTCGTGATC CGCTTGTCTC AGCCTCCCAA AGTGTGGGA TTACAGGCGT
29251 GAACTACTGC ACCTGGACAT TTTTTTTTTT TTTTTAACTT GAAAGAACAG
29301 CTAACAGACA GATTAGAACAA GAATTGGCTA TTTGACAGAT TTTCTCAGAT
29351 GAACTGTGAT AGTCATTCA AGGGAAAGTAG CTGCAAGCAT TTGTTGGCTG
29401 AAATAAAATT TAAGTTTATC ATGGAAAATT AGAATTGAA AAAACTTAA
29451 GTTTACCACT TGACAGTATC CTAATACAT ATGACTTTT TGATGAGTGC
29501 CGATATTAAT GAAGGTTATT TAAATAAT TAAATAATGT ATAATTCTTT
29551 TTATATAACA GTTAAAATA AACCATGAG TACTAGAATA AAACATAGGT
29601 GGCTCTTAA TCTTGGTTTG TGAAGTTT TTTTAAAATA AGAAAAAAAGC
29651 AAGAAATCAC TGCTAAATT GACTATTAAA ATTAATTAT CACAGGCACA
29701 AAAATGTTAG AAAACTAATG GCAATAGCAA ATATATATAT ATGAGGATTG
29751 GTATTCTCAA CATATAAAGC ACATTTGCAC ATCAACAAGA AAAGAATATT
29801 TCTCCTAATG GAAATAGTGG CAAATACATG AGCAGTCAGT TGAAAAAAGA
29851 AGTAATACAA ATTGCTGGCT GGGTGTGGGT GGGGTCAACGC CTGTAATCCC
29901 AGCATTAGA GGCTGAGGCT GGCGGATCAT CTGAGGTCAAG GAGTCGAGA
29951 CCAGCCTGAC CAACATGGAG AAACCTGTC TCTACTAAA ATACAAAATT
30001 AGCCGGATGT GGTGGCGCAT GCCTGTAATC CCAGCTACTT GGGAGGCTGA
30051 GGCAGGAGAA TTGCTTGAAC CCAGGAGGCG GAGGTTGTGG TGAGTCGAGA
30101 TCGCACCAT TGCACCTCAGC CTGGGAAACA AGAGCGAAAC TCCATCTCAA
30151 AAAAAAAAAA AAAAAAAAAA AAAAGGAAGT AATACAAATT GCCAATAAAAT
30201 ATGGAAAAAA AAAAGGCTC AACTTTATTT CTAATTAAAG GCCTTTAAGT
30251 TAAACTTAGG TGTCATTTAA TTTTTTATTAA ATTGGCAAAT ATTAAAATT
30301 AGCATAATTG TTAAGCAACT CTCGGTAGGT GGGAAAGAATC TAGCTGTAGC
30351 CTCAGGTGTT TGTGCCTCAA GGAAAACCT CTCTGGGATG TCCATTGCTT
30401 GAAGTCAAAG GTTTCCAAT AATACCTGGA AACTATTTT AAAATGCTGA
30451 TCCCCATACC CTCAAAATAT TAATAGAGAC AATCGTGAGG ACTATAATAA
30501 AGAAATGTGC AATAAGCTCT GGGGGCACAG AGGGAAGAAT CTATTGGCTG
30551 AGGAGTTGAA GAAATTGTTT GGACACTCAG TATTGCTGAA GCTCAAAACT
30601 GAAGGATGAA TAAATGCCAC ATGACCTTGG GGCTGGGGAG TAAGTAGGGT
30651 TATGCAGAGA GAGATAACTG AGGTTTTGG GCAGACGAAT AGTAACGGCT
30701 CAGGCATGGG AGTAAAGGTC ATTTAGAGAT TTACAAGAAT TCAGCATTTC
30751 TTTCCTTTTC TTTTTTTTTT TTGAGATGGA GTCTAGCTCT GTCATCCAGG
30801 CTGGAGTACA GTGGCATGAT CTCAGCTCAC TATAACTCCC ACCTCCGGG
30851 TTCAAGTGTAT TCTCATGCCT CAGCCTCCCG AGTAGCTGGT ATTACAGGCG
30901 TGTACTACTG TGCCCTGGCTA ATTTTGTAT TTTTAGTAGA CATGGGGTTT
30951 CACCATGTTG GTCAGGCTGG TCTCCAACCTG CTGAGCTCAA GTGATATGTG
31001 CACCTCTGCT CCCCAAAGTG CTGGGATTAC AGGCGTGAGC CACTGTACCC
31051 GGCAAGAAAT TCAGTATTTC TATCCAAGTA CCTGGGGGAT AGATGTGCTA
31101 CATGAATATT TATTGCAATT ATTGTTGTTCT CTGCATTTTT TTTTTTTTTT
31151 TTGGTTTGAG ATGGAGTCTC GCTCTGTGCG CCAGGCTGGA GTGAGTCGT
31201 GCAATCTCGG CTCACTGCAG CCTCCACCTC ATGGGTTCAA GCGATTCTCC
31251 ATCTGGTCT CTCAGCTAGC TAGGTTTACA GGCAGTGAGC ATCACACCCA
31301 CTAATTTCGT GTATTTCAG TAGAGACAGG GTTTCACCAT GTTGGCCAGG
31351 CTGGTCTTGA ACTCCTGATC TAAAGTGAGC CTCCACCTT GGCCCTCCCAA
31401 AGTGCTGGGA TTACATATGT GAGCCACTGC GCCTGGCCTC TATATACTTC
31451 TATAGTACCT GATACTTATT AGGCACCTAA TTACAACATA ACTTTTTTT

31501 TTTTTTTTTT TTTTGAGACA GAGACATGCC TTGTCGCCTG GGCTGGAGTG
31551 CAGTGGCACA GTCTCGGCTC ACTGCAACCT TCACCTCCCG GGTCAAGTG
31601 ATTCCTCCTTC CTCAGCCTCC CGGGTAGCTG GGATTACAGG CGCCC GCCAC
31651 CACGTCCAGC TAATTTTTTG TATTTTTAAT AGAGATGAGG TTCACCAC
31701 TTGGCCAGGC TGATCTAAA CTCCCTGACCT TGTGATCCAC TCACCTTGGC
31751 CTCCCAAAGT GCTGGTATTA CAGGGTGTGAG CCATCATGCC CGGCCATAT
31801 TTCTAAAAAC ATTTCTTAT AAAATGACAT TGCCATTATC AACCTGCAAA
31851 ATACATTTCC ATTTGGTTGT TTTCTTGCTT AGTCTTTAA TCTAGAGTTT
31901 TATACTTAT CTTTTTTATT TATATAATT TTATGTCATT GACTTTTGCG
31951 AGAAACTGAA GCACCTGTCC TGTAGATTGT CCAATATTCT AGATTTGTCA
32001 TTTTGTTC TTGTGATGTC CTTATGCTTA TTTGTTGTC CCTCTTTCTG
32051 TAATTAGAAG ACCTAGAACT GCACTATCCT TAGAGTAGCT ACTAGCTCTA
32101 TGTAGCTATT TAAATTTAAA TTAATTAAAA TTGAAAAAGT TTGGTGGCTC
32151 ACACCTGTAA TCCCAGCACT TTGGGAGGCC AAGGTGGAG GATTGCTTGA
32201 GTGCAGGAGT TCAAGGCTTC AGTAAGCTAC GATTGTAAC TAGCCTGGGA
32251 GACATCAAGA CCCTGTCCTT TTAAGGGGAA AAAATAATTG AAAAAATCAA
32301 AAACCTAGTT TCCTTGTTC ACAAGCTGCA TAGGGCTAAT GGCTACCATA
32351 TTGGCTAGCA CAGCTTATAG AACCTTTCCA TTGTACAGA AAGTTCTGTT
32401 TGGCAGTGCC GTTCTCATT GACCTGATTG GATTAAGGTC CATCTTGT
32451 GACAGACTAC TTCTTAGGTG GTGCTTTGTG GTTCATATGA TGATAGCCTG
32501 GTCTGTTCAT TCATATATCT TTTCACGAGA AATATTTTTA TTCCATTCTG
32551 AATAAAATT CATGGCAGGT ACTTGCAAGA AGCAGTTATA ATTTAAAGT
32601 TTAACATTAG GTTAAAAAAAT TGACAGGAAA CATAATATTCA CAGGTAAAAC
32651 TTGTACACAA ATGTTCATGG CAGCATTATT CATAATAGCC AAGAAGTGG
32701 ACAACCCAA ATCAATTAT GAATGGATAA ATGTTGTAT ATTTGTAGTA
32751 CATGTAATAT TATTAGGCCA ATAAAATGGG CCAGGCATGG TGGCTCACAC
32801 CTGTAATCCC AGCACTTGTG GAGGCTCAGG CAGGGGATC ACTAGAGGTC
32851 AGGAGTTGAGT GACCAGCCTG ACCATCATCA CGAAACCTG TCTCTACTAA
32901 ACGTACAAA ATTAGGCAGG CGTGGTGTG CACGCCTGTA GTCCCTACTA
32951 CTCAGGTGGC TGAGTCATGA GGATTGCTT GACCCCCGGGAGACAGGTT
33001 GCAGTGAGCT GAGATCATGA CACTGCACTC CAGCATGGC AACAGAGCAA
33051 CATCCTGCCT CAAAAAAAGG AAAAAAAAGTAA AAAAGAAGTA CTGTTACATG
33101 GTACAACATG GATGAACTT GAAAACATTC TGCTAAATGA AGGAAGACAG
33151 ACACAGAGGG CCACATATT TATGATTCCA TTATACGAA ATGTCACAAA
33201 TTGGCAAATC TAAAGAGAAA GTAGATTAGT GGTCGCAGG GAGTGAAGAC
33251 GGGTTCTTTC TGGAGTGAAG AAAATGTCCT GGAATTCTGTG GTTGTAGTT
33301 GCAACCTTGT GAATGTATAA GGACCACTGA ATTGTCCACT TCAAAAGGGT
33351 GACTTTATG TTATGTCAT TATATCTAA AAAAAATCA TAATTAGGAA
33401 GCAAGATTGA CTTCTAAGAA AAAGCGGAGT GAAATTGTTG TTTGTGGTG
33451 AATAAAATTGG GTGGGGGGGT CGCAAGAGTT TTGCTGATTA GTGATTAGAA
33501 AAATTATTCA TAATCATTGA AAATATAAAA TATTTTTCTA TATGATGTAT
33551 GTAAAGAATT TGGCAAGAGA TGATGTTGG AAAAAATAAA GAATGGCTAT
33601 TGTAGAGATC TTAAGGAAAG AAACATACAGT TAAGTAGTGC TTTGTAATCA
33651 GAATATGAAG TAAGTACTGA AAGTGGATGG AGTGGCTGTT GTCAGCATGT
33701 TATACTTAT ACATTTCATT CATAAATTTG GACTGTAGAT AAAAGTAAAC
33751 TTTTTTTTA TTTACTCTTG ACAACAGTT TTTTTTTTC CACTTAGACT
33801 TGCACTGCT CCACTGAACA ATACATTAA TTGTTAATTA TTTCCCCCTT
33851 CAGGATGATA CATATACAGA AAGCTACATC AGCACAATTG GTGGGATTT
33901 CAAATAAGA ACTATAGAGT TAGACGGGAA ACAATCAAG CTTCAAATAG
33951 TAAGTGACTT GGCTAGTAAT TTTTTGAAA TTTATTGTTG TAAATTTGTA
34001 ATGTATTGTT ATTTGTATA TATTTACTAT GCTAACAAA TTGAATGTAA
34051 AATGTCCTAA GATTGATGTA CTTAAGATAG AATGGTAGAA TAAGAATTAC
34101 TTAGATTAAA AATAATATT TCAAGATTAC TTAAGCCTCA TTGAATTITC
34151 TGTTCATGAA GCAGGAAAC TCATGTTTA AGTCAAACCTT GGTCTCTCATC
34201 TTTTTCTTT ATCAGTGGAA ATCTAAGTTC AAGTTACCT TGTCTACAC
34251 TGCACATTAA ATAGACCTT TTGTTTGTC TTTTACTGTG CTAAGTGCAT
34301 GGAACATTAA AGGAACCTTA GGAAGAGATT CTTCATATGT GGCTCAGTTG
34351 AAGAGAAGTA CTTATGTAGT TCTAAGTATT TTTATTAGAT AGTGTGCACC
34401 AACCTGTAG AAACACAGAA TTGTTGGAA AAAAGGAACCT TAGTTTTGT
34451 AACATGTTCA TTTTACTGCT CAAAAAAACG AATGCTGAAA GATTAAATGA
34501 CTTGCCTACA GTTACTGGTA GAACCAAGTG ACCGAAGCTC TGTCTTCAT
34551 ATTTGTGTC TGTGTGCCAT CCTATCCCCC TTATCCATCT TTACACCCCC
34601 AGCCCCCAAT TAAATATAGG CAATTATAAT AGTCAGTTG TGCCTCTTCA

34651 GTATGGGTCT GAGTCCTGTC AGTGTGGCA TATCTGTGGT CTTTTAAAAA
34701 ATAATCTCT CAGTATTTT CAGAGTAGGC TATTAGCAAG AAGTAGGCTA
34751 TAAACACAGG AAACCAGTGA CTGCCCTTT TCATGAACT GATGACACAT
34801 GGAATTGGAA GGAGTCCTGC ATTAGGAGTC AGAAGACTTA GATTTGTTGT
34851 CTTGGTTCTA GTATTTACCT GTTAGAGAAAT CATGGTTTG TGCTCTGGG
34901 GAAAAGGCG AAGTAACCCCT GAGACCCAGT TTCCCTTCTA AAATGTGTGT
34951 GATGACACCT GATTTACTAA TTTATAAGCT AGTTGTGAGA ACCAACTGTA
35001 ATAGCTTGT CTATGTGACA ATACGTGTGA AAGCCCTTG TAAACTTTG
35051 GGCACCATAT AGATACTACT TATGATATGA CATGCCAGA TAAATGGGTG
35101 TTTGATAGGT TAAGTTGCTC CCTTTTCTTA CATGACTCTG ATGAGGAAAA
35151 GAAGGTATGT TAACAAAAGA TAGGTGGCTG TGGATATTGA TATAAGTAAA
35201 CACACTTGAT GTGTCAAATT AGGACTTGCA AGGATTAGT TTTCAGAAAT
35251 AGCTGAAAT ACTTCAATC AGTGAACAAA TTACCCCTCA TATTTTTCC
35301 CACGATATAA GTACAGTCTC AACCTTTAT TTGGCACCAT AAAGAGCACA
35351 TAAAGATCTA CCCAAAATG TACTTTAAAG CACTGGTATG GAATAATTGT
35401 ATTATGTGTG ATCATTGGTG TTTATAAGAT TTGGGTGTGT ATTCTGTGT
35451 GAAACATTCA TATTTGTTA CTTTCCCTGTG GCTGGAAGGG ATCTTATAGG
35501 ACAGTGTCT TCATCTTGT CTGTCTTCA TCTTAATAG GAATTCTTT
35551 TCCATGCCCTG AAGGCCCTCAT TTTGAACATT TTGTTTGTGT GTTTTTTAT
35601 TTTTGAGAT ACAGTATTGC TCTGTCTCCC AGGCTGGAGT GCAGTGGCGC
35651 GATTTGAGCT CACTGCAACC TCCGCCTCCT GGGTTCAAGT GATTCTCCG
35701 CCTCAGCCCTC CCTAATAGCT GGGATTACAT GTGTGTACCA CCATGCCCG
35751 ACAATTTTT TTTTTTGAG ATGGAGCCTT GCTTTGTCGC CCAGGCTGGA
35801 GTGCCAGTGG TGCAATCTTG GCTCGCTGCA GCCTCCGCCT CCCAGGTTCA
35851 AGCAGTTCTC TTGCCTCAGC CTCTGAGTA GCTGGGATTA CAGGCGTGC
35901 CCACACACC CTGCTAATT TTTGTATTT TAGTAGAGAC AGAGTTTAC
35951 CATGTTGGTT AGGCTGGTCT CGAACCTCTG ACCTCTGTAT CTGCCCTGACT
36001 CGGCTTCCCA AAGTGTGGG ATTACAGGC TGAGCCACTG TGCCCAGCCT
36051 TCCGATAATT TTTGTATTT TCGTAGAGAT GGGATTTCGC CATGTTGGCC
36101 AGGCTGGCT CAAACTCCCT ACCTCAAGTG ATCCACCCGT CTTGGCCTCC
36151 CAAAGTGTG CGATTACAGG CGTAGGCCAC CACGCCCTGGG TTTTGAACA
36201 TTTTAAGAA GCTTACATT TTTCGAAAT AGCTAGTTCC ATTTTACACA
36251 TAACTTCAGC TAGGCATGTT GCCTCATGCC TGTAATCCCA GCACCTTGGG
36301 AGGCCGAGGT CAGAGACTCA CTTGAGGCCA GGAGTCAAACA TAGCTCCTGT
36351 GACCAGCTG GTCAACATAG AGACTCTATC TCTACCAAAA AAAAAAAA
36401 AAAAAAGTAAC CAGGTGTGGT GGTCCATGCC TGTTAGTCCTA GCTCCCCAGG
36451 AGACTGAGGT GGGAGGAATG TTTGAGGCCA GGACTTCAG GCTGCAGTGA
36501 GGCAAGATG CACCATTGCA CCCCAGCTT GGGGACAGAG TGAGAGACCC
36551 TGTCTAAAA ACAAAATAAG GCTGGGCGCA GTGGCTGTCC GGGCGTCGTG
36601 GTTCACGCTT ATAGTCTTAG CACTTTGGGA GGCCAAGGTG GCCAGATTGC
36651 CTGAGCTCAG GAGGTCTAAG ACCAGCCTGA GCAACATGGC GAAACCTCAT
36701 CTTTGCAAAA CATAAGAAA AAAACAAAAA AAACCACAAA ACCTCTAGTT
36751 GCCAGTTATT TTTTTTATTT ATTCTCTAGTG ATTCTCTTT TTTTCTTTT
36801 TCTGAGACAA AAATTCACT TTGTCCTCCCT CGCTAGAGTG CAGCGGTCA
36851 CTCACTACAT GATTCTTTA GAGACATGTT AATTCTTTAT ATTGAGCTGA
36901 AGCCTGTTTC TTTTACTCT GTCTCTCTT ATTCTCCCGC CTTGTAGAGC
36951 TGCCCTGAATC AGATTAATC CTCTTTTATT GGCAAGCCTG CCCTCAGAT
37001 TGATCTTATC ACAACCTTC TTCTACCTCT GAAGTCCTCA TTCTTCCGT
37051 TAATGATATT TTCAGAACCT TGTGAATT TGGTTATTCT TACATTTTAT
37101 AAATGCCTT TATTAATTT GATTTCTTAA ATCAAGTATG AGATATAACA
37151 CATGAGGTAATCCTGTCTT GATTTGGGAGC CTGAATGAAT TTCTCTCTTG
37201 AACTTCAAGG GCTCATGGCC CTTTCTTATT ATTAATCAAA GACAACCAATT
37251 TGGTGTTCAGT GTAGCTATAT TTTTCTAGT TTGGGTCTTA AGGTTTTTGA
37301 TTTGCTTGTGTT TTTTCTTTT TCTTTTTTTT TTTTTTGAGA CGGAGTTFCG
37351 CTCTTGTGTC CCAGACTGGG AGTGCAATGG CGTGATCTCG GCTCACTGCA
37401 ACCTCCGCCT CCCAGGTTCA AGCGATTCTT CTGCCCTCAGC CTCCCTAGTA
37451 GCAGGGATTA CAGGCATGTG CCACCAAGCC GGGCTAATT TGTATTTTGA
37501 GTAGAGATGG GGTTCTCTA TGTTGGTCAC GCTGGTCTCG AACTCCCGAC
37551 CTCAGGTGAT CGGCCCTGCC TGGCCTCCCA AAGTGTGGG ATTACAGTCG
37601 TGAGCCACGG CGCCTGGCCG ATTIGCTTGT TTTTAATCAA AATAGGGGCC
37651 TTGGCCAGGT CGAGTGTTC ACCCTGTAA TCCCAGTACT TTGGGAGGCT
37701 GAGGCAGGCA GATCTCTGA GTTCAGGAGT TCAAGACAG TATGGGCAAC
37751 ATGGTGAAC CCTGCTCTA CCAAAACAC AAAATTCAAGC CAGGCATGGT

37801 GGTGTGTCCTG TGTAGTTCAA GGTACTCAGG AGGCTGAGGT GGGAGGATTG
37851 CTTGAGCCCC GAGATGGAGG TTGCGGTGAG CCAAGATTGT GCCATTGCA
37901 CTCTAGCCCTG GGCAACAGAG CGAGACCTTG TTTCAAAAAA AAAAAAGAAG
37951 AGGGTCTCAC TTTACACTTC TGTGACTGGT GTTTTAAAAAA TCTAAACACA
38001 GGCCGGGCAC GGTGCTCAC GCCTGTAATC CCAGCACTTT GGGAGGCAGA
38051 GGCACGCAGA TCACAAGGTC AGGAGTTCTG GACCAGCCTG GCCAGCATGG
38101 TGAAGCCCAT CTCTACTAAA AATACAAAAA AATTAGCTGG GCATGGTGGC
38151 AGGTGCCTGT AATCCCAGCT ACTTGGGAGG CTGAGACAGG GGAATCACTT
38201 GAACCCAGGA GGCGGAGATT GCAGTGAGGCC AAGATTGCGC CATTGCACTC
38251 CAGCCTGGTG ACAGAGCGAG ACTCCGCTCTG AAAAAAAA AAAAAAATCT
38301 AAACACAAGA TTTTACTTTT AATCCTATCA TTTCTCTTGT CTTGGCTTCA
38351 GTAATCCTTC AAGTTTCTA GGTCTTTCA AAATCTTGAT TCTGTTGATT
38401 TATATTTAA TTATCTTTTCTT CTTTCAGCTT TTCCCTGTTCA GGTGTCACAT
38451 CTGGGTCTTT ATCTGAGTTT TATTAGATTA TAAAACATTC AGCAAGATAG
38501 GGCAGGTACT GAGTCAGTT GTACACCAGT GAAGGCCTCT TTCTGTGATT
38551 GTTCATTCTAT GAGGCTTTAT GAAAATGTCT ACATTACACC AGGACTTGG
38601 AGGTTACAGA GATGAATAAA ACATAGTCCA TTAGGAGGCA GACAATGGGA
38651 GAGACAAACAA TGGGAAAAAG TTACTCTGAT TATGAGGAGT AATGAGAATT
38701 ACATATGAAG GAAAGTATTG TTAGTACTGT TAGGATTAG TGTCAGGAAA
38751 GTTTTCAGAG TAGCAAGGAA ACATCAGAAA TTTTACTCTT TCTGCCAGGC
38801 ATGGTGCGAT TATTATTCTG TTCTCACACT GCCACAAGGA ACTGACCAAA
38851 ACTGGGTGAT TTATTAAAAAA AAAGGTTAA TTGACTCTATA GTTCTGCATG
38901 GCTGAGGAGG CCTCAGGAAA CTTAAGTGTGG CAGAAAGGGA AGCAGGCACG
38951 TCTTACATGG CAGGAGGCGA GAGAGTGTGA AGGAAGTGAA GGGGAAAGAG
39001 CCCCTTATGA GACCATCAGA TCTTGTGAGA ATTCACTCAC TATCACTCGA
39051 ATGGGGAAA CGTCGCTCAT ATCCAATCA CTTCTCCATA ATCCAATCAC
39101 TTCCCTCAGT GATTACAATC TGAGATGAGA TTGGGGTGGG GACACAGAGC
39151 CAAACCATAT CAGTCGCTGT AGTCCCGAGTT ACTTGGAGGC TGAGGCAGGA
39201 GGAACACTTG AGCCCAGGAG TTCAAGATCT GCCTGGCAA CATAGCAATA
39251 CCTCATTCTT GGATAAAAAG GAAATTTCAC TTTTGGGTG CCATTGCTTA
39301 GTTAAATCAG CTGTAACCTTC TTGTTGACTT TTAGTCAAAA AACAAATTTC
39351 CCTTCTATCT TTGTAAGAAGA GGTTGGTAGAG CAAGGAAGAA AAGGAAACTT
39401 GCTTTATTGA GCAGCTTCTA TAGTCAGGCA CATTTCACAA ACATTAGTTC
39451 ATTAAACCC CTTAGCTGT TGTACAAGGT GAATGCTATC TAGCATTAC
39501 AGATGAAGAA ACTGTTAGGT GACTCTCCCT AATATTAAT AACCAAGAAC
39551 CTGGATTGAGA TGTTTTGAG TCAGGGTAGC TTGATCCTCG AGTTCATGCT
39601 TCCTCCAAGG ATACACTGAA AGACTTTGAG CCTCTTTTTT TTTTTTCTC
39651 TTTTTTGAG ACAGGATCTG GCTCTCTTGC CCAGAGTGCA GTGGTGTGAT
39701 CTCAGCTCAC TGCAACCTCT GCCTCTGGG CTCAAGCGAT TCTGCCTCAG
39751 CCTCTCGAGT AGCTGGGACC ACAGGCCAC CCCAGCATAC TTGGCTAATT
39801 TTTGGATTTC TAGTAGAGAC AGGTTTCA CATGTTGGTC AGGCTGGTCT
39851 CGAACCTCTG AGCTCGTAAT CGGCCGCTC CGGCCCCACA AAGTGTGGG
39901 ATTACAGGGC TGAGGCCACCG ACCCACTTCA AACAGTTTT TAAAACCCAG
39951 AACTATAATG CAATAATGTT AGCATTGTT TTGGGAGTTT GAGCCTAAAT
40001 GGTTGAAGTG CAGTAAATTG TTCTTAAAT ACGTTTATG AAAGTATTG
40051 GAGTCTCTTC CTTACATTTC TTCTCTAGC ATGAAGACAA CACCTAGCCA
40101 GGCAATGGTGG CTCATGCCAG TAATGCCAGC ACTTTGGGAG AATGAGTTAG
40151 GATAATTGCT TGAGTCCAGG AATTTGAGAC CAGCTGGGCA AATGTAGCGA
40201 GACTCTGTCT CTACAAAAAA GAAAAAAATTA GCCGGGTGTG GTGGCATGTG
40251 CCTGTAGTCC CAGCTACTCA GGAGGCTCAG GTGGAAGGAT TGCTTGAGGT
40301 GGGAGGTGAGA GGCTGCAGCC AGCCATGATC ATGCCACTGT ACTCAGCCTG
40351 GATGACAGAA TGAGACGCTG TTGAGAGGG GAAAAAAAAG ACACCTGCTT
40401 GGGATGATTA AAGTTCTGTC TTGACTGGTA GTTATTGAA TTAGGTCCCT
40451 CCAGTGTCTT TAATCATGGT AGAATGTGCT AGCAAGTGAG TTTGCTTTAC
40501 ATGGAAGAGT TCTGTGTTCA AGGGCTTTCG GCCAGTGGCA TTCCTAAACA
40551 CAGTGTAAA GGCGGTAGGG AATGTGAAAAA GTATGACATA GTTCTGCTC
40601 TCAACAGCTT GTAATTTCAG TATTATTATC GTAAGCTCAA TTGTTAGGTAC
40651 TACTTCTTTT CTGGACTTTC AGGGTGTCTTAC TCCGTGCAA TTTAGTGGTA
40701 TGAGTTGAGG ACTAATGTTT CTATATCACA TCCTGATAAT CTCCACAGTT
40751 ATGAAAACCA AACTATTTC CTCCTCCT ACACTTTCC CCAACTTTAT
40801 TTTAATGGAA TTGTTGGAT TTCTTGATTG TTTTGTAAATA GTGGGACACA
40851 GCAGGCCAGG AAAGATTTCG AACAAATCACC TCCAGTTATT ACAGAGGAGC
40901 CCATGGCCTAC ATAGTTGTGT ATGATGTGAC AGATCAGGTA AGTTCCAAGA

40951 GGAGATTGTG TTACAGTGAC CAAGTAGGAA GCCATTATTT GATTAATGTC
41001 AGATTCATTT ACTACTTCAT ATATAAGCCA TCAGTATTAA TTTTATGGCA
41051 GAAAACCTTG TCCACTCTCA AATATAAAATG TGAATCACTT AAAAGACATT
41101 TGTTTCTCG TAATAAATAA AAGATTAGTA ATTAGTTTA CGTTTGCTTT
41151 CAAGGGATTC TGTTGTATT TATTGTCAAC TAAATAACTT TGATCAAATA
41201 GCAAGACTC TAACATATAG GCAAGAGTT GTAGGGAAATC GTGAGTTGCT
41251 TGGCTTATAC TGTGTTCTTG GTGTTAAGTA TTAACAGGAA TATGGCCTGG
41301 TAATTAGAAC TTGTCATCA GAATTGCCAA AAGTGGGATT CGGGGGTCTC
41351 TGCCATATGGA GGATGTGGTT CAGAAATAAA GAATTGAAT AGGATAAGCT
41401 GTAGGAGGAT CTTAGTATGA GAATGAGTAT CTGAAGATTA GCTGTGAGAG
41451 AGGCCAGAGC GATGGAGGGA ACAATGTGGG ACAGTGTGAA GCATGTGATC
41501 CAGGGCCAT AACTTTTTT GTTACTATT TTTTAATCA GAAACTTAGA
41551 TTCACTGTC CTTTCTATCA AAGAAAAGGA CAAAAGATAA ACGTTCAAAA
41601 TTGGAATTAA TTTTCTTTT GGCAAATGTT AAATCTCACC TCTAATGAGA
41651 AATCATAGCT AATTAGGAGA TAACTTACAT GTAAGCATT AGATTCACTG
41701 CCATTAGAAC TGCTGGTGG GTGATATCTG CAGGAGAAAA AAATGATGCT
41751 AGTTTAAAAA ATCTCTACTA TTACCGTGAAT ATATTTTAA ATGAAAACCT
41801 TCGTCCTCTA AATATGACTG TGGAAAAGAA AATGAGTATA TTTAATAACA
41851 TCTTTGACA TCTCTAGTAG TAACAGTAGG TCATCTTATT CATAAACCAA
41901 AATTTTACCA AATTCAGGC CAGGCGCAGT GGCTCATGCC TGTAATCCCA
41951 GAACTTTGGG AGGCCGAGGC GGGCGGATCA CCTGAGGTCA GGAGTTAGAG
42001 ACTAGCCTCG CCAACATGGC AAAATCCCCT CTCTAGTAAA AATACAAAAAA
42051 TTAGCCAGGC GTGGGGGCCG GTGCGCTGAA TCCTAGCCAC TTGGGAGGCT
42101 GAGACAGGAG AATCGCTTGA ACCCAGCGGG CAGAGGTTGC AGTGAGCCGA
42151 GATCGCGCCA TTGCACTCCA GCCTGGATGA CAGAACAAAGA CTTTGTCTCA
42201 AAAAAAAAAA AAAAAAAAAA AAAAAAAATTA ATCAAAATTTC AAAACCCAGGT
42251 TTTGTAGTAC ATTTAAATTG CATATTCCAA AGCAGTTGGG TTGCGCTGCG
42301 TTGCAAGTTA ATATTAAGCT ATACTTCCCT TTCAAATAAG GTATTTTCA
42351 CGTAAAGCT GTAAATTCTA GTTGTCAATT GTTTAGATAT TTATAGTCAT
42401 TTTAATATAT CTGTTTACGG CCAGCTGCAA TGGCTAACAC CTGTAAACTC
42451 AGCACTTTT GAGGCCAAGG TGGCCGATT GAGCTCAGGA GTTCGAGACC
42501 AGCCTGGCA ACATAGTGA ACTCCATCTA TACAAAAAAAT CCAAAAAAAA
42551 AAAGACAGGT GTGGTGGCAT GTGCGCTGTAG TCCCAGCTAT CCCGGAGGCG
42601 GAGGCCGGAG GATGCGTTGA GCTTGGGAGG TCGAGGGTGC AGTGAGCTGT
42651 GATTGTGCCA CTGCACTCCG GCCTAGGTGA CAGAGCAAGA CCCGTCTCA
42701 AAAAAAAAAA TCTCTTCACT CCTTAGCAGT GGTTATTTTG TAGCTAGAGT
42751 TGTCTCACTA GCTCTTGTGTT ATTTGTCTGT TAGGTCAAGGA ACGATGTTTC
42801 TGTTTATTCC AGAACTATAT TATCGAACTA TATTATCAGT CTTCAAAATG
42851 TCTTTTTAGG AGTCCTCAA TAATGTTAAA CAGTGGCTGC AGGAAATAGA
42901 TCGTTATGCC AGTAAAAATG TCAACAAATT GTTGGTAGGG AACAAATGTG
42951 ATCTGACCAC AAAGAAAGTA GTAGACTACA CAACAGCGAA GGTATGTTA
43001 AAGTTTAATT TTCATACTGA ATTTGAAGGT GTTGAATTAT GTATGGGTT
43051 TCGAGTAACA GTAAGGCCAC AGCCTTTAA AAATATGTGC ACTAGAAATAC
43101 TGTGACAGTG ACAATTGTG TAGCATCTGT TTGGATCCAA TGAACCTAGT
43151 TCCTCACGCT CCATTATGGA TGGTAGAAAT GCAGTAAGAA TTAGTAAAAA
43201 AGATTTTCA GTGTTAATTG TGCCTCATTA TTCTCTTAGG AATTTGCTGA
43251 TTCCCTTGGAA ATTCCGTTTT TGGAAACCAAG TGCTAAGAAT GCAACGAATG
43301 TAGAACAGTC TTTCATGACG ATGGCAGCTG AGATTAACCAAGA GCGAATGGGT
43351 CCCGGACAA CAGCTGGTGG TGCTGAGAAG TCCAATGTGTT AAATTCAAGAG
43401 CACTCCAGTC AAGCAGTCAG GTGGAGGTG CTGCTAAAT TTGCGCTCCAT
43451 CCTTTCTCA CAGCAATGAA TTGCAATCT GAACCCAAAGT GAAAAAAACAA
43501 AATTGCGCTGA ATTGTACTGT ATGTAGCTGC ACTACAAACAG ATTCTTACCG
43551 TCTCCACAAA GGTCAAGAGAT GTAAATGGT CAAACTGAC TTTTTTTTTA
43601 TTCCCTTGAC TCAAGACAGC TAACTTCATT TTCAGAACTG TTTTAAACCT
43651 TTGTGTGCTG GTTTATAAAA TAATGTTGT AATCCTTGT GCTTCTCTGA
43701 TACCAAGACTG TTTCCCGTGG TTGGTTAGAA TATATTTGT TTGATGTTT
43751 ATATTGGCAT GTTTAGATGT CAGGTTTAGT CTTCTGAAGA TGAAGTTCA
43801 CCATTGGTCA TCAAACAGCA CAAGCAGTGT CTGTCACCTT CCATGCATAA
43851 AGTTTAGTGA GATGTTATAT GTAAGATCTG ATTTGCTAGT TCCTCTTGT
43901 AGAGTTATAA ATGGAAAGAT TACACTATCT GATTAATAGT TTCTTCATAC
43951 TCTGCATATA ATTTGTGGCT GCAGAAATATT GTAATTGTT GCACACTATG
44001 TAACAAAACA ACTGAAGATA TGTTAATAA ATATTGACT TATTGGAAGT
44051 AATATCAAAC TGTATGGTGA TAAGTATTGT TTTGATTCTT ATGGTTAAAG

FIGURE 3, page 14 of 21

44101 GGAAATAGAG CCTTGCATTA TATTCAACAC AGCCATTGT GTGTGCACAA
 44151 TGCAAACATAA GGTATTCTAG ACCTATCTTA GAGCAGCATC CAGTATTG
 44201 TTTCTAGATA ATATGCCAA TAACATGACC TAGAGGGCT TCTGTGCTGT
 44251 GTAGGGATT AACCAACTTC AGTGGTCAG GGAGCTAAA CTATATGTA
 44301 ACAAGTTTA GAATGTATGC TATCTAGCCC GTTATCTCTG ATCCTTCTCT
 44351 AAAACCATTT GAAATAGCTT CATTGATCAA CATTCTATAA ATGCATCTGT
 44401 GGTAGAGGTA GAAAGCAGCA CCTTTCTAA TTGGCAAATG ATCAGACTAA
 44451 TGTGTGCTAA TGTTTTCTT CCATGCTTTC AGTCAGATT AACTATTTA
 44501 TCCTCCACAG TTGCTTAACG TGCTGTTGGA GGAGGGTTA AGCATTAAGA
 44551 TAGGAAGCAG GAAATTGAT TGCTCTAAAT TTAGAAATTA TATCCCTAA
 44601 AATTAAACAA TGAATACTGG GTGGTAATGA TAATTGAGGC AAATGTATTT
 44651 ATTTGGTGA CATTGGCAT ATATGAAGAT TTTCTGAAAT AGGACCTTCA
 44701 AGATCCTAGG GGGTTTTGTT TGTTTTTAA TTGTGAGGAA TAAAAAAATCT
 44751 TCTGCCACA CTGGCATTAAAGGTGACTG AGGTCAAACG TTGTTCCCTT
 44801 AGGTTGAAAT AGCAGCCAAA ACATTCTCA CGCAGGGCT TGGGATATGG
 44851 CTGCTGGCAA CACATTGTGTT TGTTGGCTCC TTAATTAAAT GATAAAATTT
 44901 AAGCTAAACAA CAAGCCAAA ATGAATAGGT TTTTTAAATT TTTATTTTC
 44951 ACTAAACAGG CAATTGAAAT ACATGGTACA AAAATAAGTG GTAAGATAAT
 45001 TGTAATGAA AATGGACAGA ATATTCAATT TTCCATCTAT GAAAATTTCA
 45051 CAATAAAAT CATAGTTAC TTGTAITAT AGGCGTGCTT GGTGGATCTA
 45101 TTCATCCTCA CATAAGGCAA CTGACAAATT CCTGAAGTTA CCAATAGTTA
 45151 TTTTGGTGAAG GATCTTTAAT GCTTCAGAAG TTTTGTGTTT GCCTTAATAC
 45201 AGTATAAAAGG GGGAAAGAGT TCAGAAAAGT TTTCTAAAG TAGCTAAATG
 45251 ACACAAACAA AATGTCAGA TACTGTGATG CCATGCCGTG CACTTCATT
 45301 TTACACAGTA AAAGTTGTTT AAATTGTCAG CTTATTCTTG GTGAGTTAGC
 45351 GGAAACATTA CATGAACTTA AGATGAGCAT ATTTACAGAC TTAAGTTG
 45401 AAAATTCCAG CGTTCTTTCC CCCAIGGCCAG TAAAGATTTGG GATTTACAAC
 45451 AAATTTCAGC ATGCCCTTAAG ATTGCTTCT ATGTATAACGC CAATAATGT
 45501 GGTTCTGAA AAAATATATA CCCCTTTATA CCCCCATTAA CAAGTACAAA
 45551 CGGTTCAAAG CTACTACAGG TTTTAATAAT CTGTTCACTT AGTAAAGGGA
 45601 ATTACCAACT GTTCTAAATA TAAGGTGCTG CCATAAAATTAA GTTTACATAG
 45651 TGAAGAAGAG TGTTCTAAA TCTAACGAGC TGACACTCT GTGAAATCCT
 45701 TTCAGAATGA TAGTCATTGT GGTCTGAGCA GTAATTTCCT ATTCTTCGAC
 45751 CTTGGATTGA ATTCCCTTA GCCTACATCT TGCCCTTCCA GCATATCTTA
 45801 CCTCAAACCT TCTTGTGTT CCATTCCAC CTAAGCTTCA AAATAGCCCT
 45851 GTGTTGACCGT CGTCTTCCAT TTGCTGAGCT TACCTATGGA TCTCCAAGAA
 45901 CCCAGATCTT GAAACTGCTG ATCCAGCTTT CAGTATCATC ACTTCCCTGT
 45951 GGATTTAACT TCCATTAATT TTAACGGACT ACTAAGTTAT TCCAGTGTGG
 46001 CATCACAGTG CAGTTAGCAA GCTCAGCTAC TTGACTCTAA TTGGCCATG

(SEQ ID NO:3)

FEATURES:

Start: 2181
 Exon: 2181-2203
 Intron: 2204-27090
 Exon: 27091-27163
 Intron: 27164-33853
 Exon: 33854-33949
 Intron: 33950-42859
 Exon: 42860-42991
 Intron: 42992-43239
 Exon: 43240-43434
 Stop: 43435

CHROMOSOME MAP POSITION:

Chromosome 2

ALLELIC VARIANTS (SNPs):

DNA Position	Major	Minor	Domain	Protein Position	Major	Minor
397	T	-	Beyond ORF(5')			
2326	A	G	Intron			

3486	C	A	Intron
6651	-	A	Intron
8190	T	-	Intron
8281	T	C	Intron
11546	A	G	Intron
11670	C	T	Intron
11688	A	G	Intron
14938	A	C	Intron
22261	G	A	Intron
22852	G	A	Intron
27253	A	C	Intron
28098	-	A	Intron
28597	G	T	Intron
31431	C	T G	Intron
35704	C	T	Intron
35728	C	T	Intron
36690	C	T	Intron
41002	G	C	Intron
41033	A	G	Intron
43161	C	T	Intron
43765	A	G	Beyond ORF(3')
44713	G	T	Beyond ORF(3')
44831	C	T	Beyond ORF(3')

Context:

DNA
Position

397

TGCTCTGCGCCAGGCTGGAGTGCAGTGGCCTCTGGCCCACGTAGCCTCCGCCTCCC
GGGTCAAGCAATTTCCTGCCTCAGCCCTCCGAGTAGCTGGGATTACAGGCACGCGCCA
CCATGCCTGGCTAATTTTGATTTAGTAGAGACAGTGTTCACCATGTTGCCAGGC
TGGTCTGAATTCCCTGACCTCGTGTCTGCTCGTTGGCCTCTCAAATTCTGAGATTA
CAGGCATGAGCCACCGAGCCTGGCCAGTTCTGAGTTTATTGAAATCAAATAAGC
[T, -]
TTTTTTTTTTTTAATGGGCTTAGAGTCCAGGGTAACGAACACTTTTGTTGCCTATT
ACTGAACCATTCAAGGGTATTCCCTGGGTGGTGACCGTGTCAATTCAAGAAACCAACATGT
TCATTTCAAGAACCAAACCTGGGTAACCTTGATAAGTTCATCAACTAAGGCCATGGCA
GAATTGAGGGCTAAGGGGTGTAATTAGTGTATGGTAGAAATAAGTGCCTCTTCTAT
ATTTGGCGTTGAGAATTAAAGTGATTCTGCAGTAAGTCTCAGGAGACAATTCTTCTT

2326

GCTGATTGTGTTCTAGGGGACGGAGTAGGGGAAGACGTTGCTCTCCGGAACAGCCTAT
CTCATTCTTCTTTCGATTACCGTGGCGCGAGAGTCAGGGCGCGCTCGGCAGCA
AGGGCGCGGTGGCGCGCGCGCAGCTCAGTGACATGTCCAGCATGAATCCGAATAGT
GAGTTCAAGGAGACCCGGTGGCTGGTGGCCAGCTGGGGATCTTAAAGGG
TCGAGGAGGGTTGGGCAGAAGTCGGGCATCGCTGGGTGAGGCGAGGGTGTGGTC
[A, G]
GGAGAGGGCTGGCGCGGGAGTCGGGCCCTATTGCTGACCGAGGGCGCCGCG
GGGAGGGTGGCGGGAGGGTGAGCGCCGGGGCTGGACGGGGTCAAGGTTAGAGGGC
CTGACTGCGGGCGGGTGTGACGAAGGCTGCCAGGGGCTGGGCGGTGTGAAGGGGT
ATCTCTCTCGGAGGCAGTGACTTTGAAGGAGGACTGTCTCAAGGGAGGGGATGGG
GTGGGAGAGCCCTCTAGAGGCCACTGTCAAGCCCTGCCGCACTCTGCCAGCTGTC

3486

CTGGGAACGGTGTGTTCACTCCCTGGTAGAGTTGTTGGCTCTCCTCAATGGCCCTT
TAAAAAATTCCCTCACAGTTACATGCATGTAAGTAATGAATAATTGAAAGAGACCGAA
TTGGTATTCCCTTCAGTGTCAAAGGCCATTGAGGGATGGGGAAAATCAGTATTGTTG
TAAAAGTTGAGTTATTGCTGGTTGGTCAATTACTGCTAGACATTTCCTCAAAGG
TCCACCCACCACTGACTGTCAATGTGTGTCACATGGCTTTGCAAAATGCTTA
[C, A]
AAGTTTGTAATAGTGTGGCTGAAGCTGAAATCTTGCACAAACAGAAACCGTAGTA
TTTATTAGAATTTCATGCTTGTAGAAGTTGAGGGTAGTGTCTGTAGTGACATTGCTG
TGTGACAGTTAAAAAAATTCTCAAGGGCTCCAAGGACAAAGTTGGTTTGAC
AGTGAACGGAGGTGAACCTGAGGTTCTTAATTAGTAGTTCTTGGTAACAATAAAGA
ACATGGATTACTGCTTATCGAGGTTAGACCTCTACTGTTCAAGGAAATTCTGAA

6651	TTTCAGCACATTAAGAAATGCTTAACATGGCCAGGCCAGTGGCTCACGCCGTAAATTCT CAGCACTTGGGAGGCCAGGGTGGCGGATCATTTGAGGTATGACCAGCCTGGCCAACA TGATGAGACACTGCCTCTACTAAAAAATACAAAATTTAGCTGGGTGTGGCAGCCT GTAATTCCAGCTACTCAGGAACCTGAGGCAGGAGAGTCAGTGAACCTGGGAGGCCAGG CTGCAGTGAGTCCAGATCATGCCACTGCACTCCAGCCTGAGGGACAGAGTGAGACTCCTC [-, A] AAAAAAAAAAAAAAAGAAAGAAATACTTAACATTATTCTCGTATTATTCTCATAAC ATTTTCATAATCCACTGGCTTCCAGTGGATTTTTAGTGTCAAGAAAATAATTGAT TGGTCATCTTAAGGAATGTGTTAAGAATAAAGCATGTCTACCTGTCAGTATAACCA GCTAACTATAGTAGGAAGAAATACTAGTAGTCAACTATAATTCTTAATGC AGAAAAAGTTAAAGTATTACCTTACCTTAAAGTATATCATGGC
8190	AGACCGGCCTGGCAATGTGGTAAACCTGCCTCTACTAAAAACACCAAATTAGCTAGG CGTGGTGGTGTGCGCTTGTAGTCCAAAGCTACTGAGGGAGCTGAGACAAGAGAACGCTT GAATCTGGAAAAAGAGGTTGCCGTGAGCCAAGATTGCCACTGCACCTCCAGCCTGGGTG ACAGAGTGAGATTCTGTCCTAAAAAATAAAATTTCCCCCTTAATCAAATT AAGTAAATGAGGGATGTTAGACAGTTAACCATCAAATATTAGTTAGTTAGTTAGTT [T, -] TTTTTAACGTTGCTTAAAGATGGAAGTGCTTCAAAATCAAATCTCCTGCCAGTTCTC TACTTGGCTTCTTTTTCTTTGAGATAGACTCTCACTTGTCACTGGAGTGCGTT GGCGTGTACTCGGCTCACTGCAACCTCCGCCCTCCAGGTTAACGTGATTCTCACCTCA GCCTCTCAAGTAGCTGGAGTACAGGTGTGCCACACACCCGCTAATTGTAGTT TTAGTAGAGACAGGGTTCACTATGTTGCCAGGCTGGCCTCAAACCTCGTACCTCGTGA
8281	CTGAGGAGGCTGAGACAAGAGAACGCTTGAATCTGGAAAAAGAGGTTGCCGTGAGCCA AGATGGCCACTGCACTCCAGCCTGGGTGACAGAGTGAGATTCTGTCCTAAAAATAAA AAATAAAAATTCCCCCTTAATCAAATTAAAGTAAATGAGGGATGTTAGACAGTT AACCATCAAATTTAGTTAGTTAGTTAACGTTAACAGTGGAAAGTGC TTCAAAATCAAATCTCCTGCCAGTTCTACTTGGCTTCTTTTTCTTTGAGA [T, C] AGAGTCTCACTTGTCACTGGAGTGCCTGGCGTGTACTCGGCTCACTGCAACCTCCGCC TTCCAGGTTAAAGTGTATTCTCCACCTCAGCCTCTCAAGTAGCTGGAGTACAGGTGT GCCACCACACCCGCTAATTGTAGTTAGTAGAGACAGGGTTCACTATGTTGCC AGGCTGGCCTCAAACCTCGTACCTCGTGTACTGCCACCCACCTCAGCCAATTGCTGGGATTA CTTGTGTGAGCCACGCCCTGGCTCTACTTGGCTTTAAAGGAATTGCTTCTGAG
11546	GTTACATTTAACCCATTATGGCGTGTAGCCATACTCACGTTACATTGATGCATCTGC TCCCTTGTGTCTATATACTCATATAACATTGCTAAAGTTAGGCAGTTCACACCA AGGCTGTTCATGAACCTCAGATTAAGAATACTGATTAGGAGATTGAAAACAGAAAAGA GAATGTTAACTATCATTATCAATATTAAAGTGTGAAATCTGAGAGTGACAAAGCTTAGC TTAAATCTGGTATCCAAACTCATTTGAGTTTTTTTTTTTTTTGAGAC [A, G] AGGTGTCGTTTGTCCCCCAGGCTGGAGTGTAGTGGTGTGATCTGGCTACTGCAACCT CCACCTCCCAGGTTCAAGTGATTCTCCTGCCCTAGCCTCTGAAGTTGCTGGGATTACAGG CTGCCACCACGCCAGCTAAACTCCTGATCTGTGATCTGCCCTGGCCTCCAAAGTG CTGGGATTACAGGTGTGCCACTGTTCCGGCCTAATTGAGTTAAAATGTGGAGTT
11670	TGTCATGAACCTCAGATTAAGAATACTGATTAGGAGATTGAAAACAGAAAAGAGAAAT GTAACTATCATTATCAATATTAAAGTGTGAAATCTGAGAGTGACAAAGCTTAGCTTTA AATCTGGTATCCAAACTCATTTGAGTTTTTTTTTTTTGAGACAAGG TGTGCTTGTCCCCCAGGCTGGAGTGTAGTGGTGTGATCTGGCTACTGCAACCTCCA CCTCCCAGGTTCAAGTGATTCTCCTGCCCTAGCCTCTGAAGTTGCTGGGATTACAGGCTG [C, T] GCCACCACGCCAGCTAATTGTATTAGTAAAGACGGAGTTCACCTTATTGGC CAGGCTGGTCTAAACTCCTGATCTGTGATCTCCGCCCTGGCCTCCAAAGTGCTGG GATTACAGGTGTGCCACTGTTCCGGCCTAATTGAGTTAAAATGTGGAGTTAAG ATGTTAGTCTTAAAGTGGGTTAGATGAAATTATAAAATAGTCAAATAGCTAAATTAT AAAAGGCCATTGAAACAATTGTGAAATATATAATGTGGATAATTATGTTAGTGTCTTAA
11688	TAAGAATCTGATTAGGAGATTGAAAACAGAAAAGAGAAATGTTAACTATCATTATCAA TATTAAGTGTGAAATCTGAGAGTGACAAAGCTTAGCTTAAATCTGGTATCCAAACT

FIGURE 3, page 17 of 21

CATTTGAGTTTTTTTTTTTTTTTTTGGAGACAAGGTGTCGTTGTCAGGCCCCAG
 GCTGGAGTGAGTGGTGTGATCTGGCTCACTGCAACCTCCACCTCCCAGGTTCAAGTGA
 TTCTCCTGCCTCAGCCTCTGAAGTTGCTGGATTACAGGCTGCCACCACGCCAGCTA
 [A, G]
 TTTTTGTTATTATAGTAAAGACGGAGTTCACCTATTGGCAGGCTGGCTCAAACTC
 CTGATCTTGTGATCCTCCGCCTCGGCCTCCAAAGTGTGGATTACAGGTGTGAGCCA
 CTGGTCCCGGCTAATTTGAGTTAAAATGTGGAGTTAACATGTTAGTCTAAAGTGG
 GTTAGATGAAATTATAAAAATAGTCAAATAGCTAAATTATAAAAAGGCCATTGAAACA
 ATTTGTGAAATATATAATGTGGATAATTATGTTAGTGTGCTTATGTGTAGATTGGTTA

14938 CATGGTAGTGTGACCTGTAGTCCCAACCACTGGGAGGCTGAGGTGGGAGGATTGCCTG
 AGGCCAGGAGTTGAGACCTGGGAGCATATGAAGACCCCTGTCTCTAAAAACTAAAAAT
 AAAAATAGCCAGGTGTTGGTGTGTTGTGCTGGTCCCAGCTACTCAAGAGGCTGAGGCA
 AGAGGGTTGCTTGAGCCAGAAGTTGGAGGCTGCCGTGAACTGTGATTGCACCACGTGAC
 TTCAGCCTGGGTGACATAGCAAGACCCGTCTCTGTGGTGGTGGTGGGTGGGGGGGGGG
 [A, C]
 AGGGATTTAAGAAGGGTTGTGAGGTATGTATTATTAATGGCTTTAACCTTAC
 CTTCACATCTGGGTTGAAATTAAATTGTATCCATTCTCAGTTCTGTCTGCTATATA
 TTTAAACTGGAGACTTAGAGGTATGGATGTCTTCTATGAAAGCAAATGAAGCAGAG
 GGCTGCCTCTCTGTGAGGGCACACTGCTGAGAGCATGTTACTGTTTATGCA
 TTGCTAGGCTTGGAGTTGTGACTTGTATGATCATAGTACTAACATTAGTTGGCA

22261 CACCCACACATACTATGTCACCGTAAGGGGGAGAAACACAGACCCCAACTCTCGA
 GGGTAGAAAATATGAGGTTATAGTAGAATTAGAACTACAAAAGCTAGAGGAAGTTCTGAA
 CTGGAAACAGTGGATAGGATTACTAGAATAATTACGAGGGTGACAATTGAAATCTTC
 ATAGGTTCTTTCTCTCTCTCTCTCTCTCTCTCTCAGTCACTGCAACCTCCGCCTCTGGGTCCA
 [G, A]
 GTGATTCTCTGCCTAGCCACCCAAAGTAGCTGGGATTACAGGCATCTGCCACCATGCTG
 AGCTAATTGTTGTTAGGTTAGAGACGGGTTTACCATGTTGGTCAGGCTG
 GTCTGAACTCTGACCTCAGTAATCACCACCTGGCCTCCAAAGTGTGGGATTA
 CAGGTGTGAGCCACCGCGCCAGCCAAATTATTGTTCTAAACTAGCGTAATTAG
 TTTTTTCACTTAAGTCAAATTATATTGTAGGATAAAAACCTAGTGTATCCAAATTTC

22852 ATCCAAATTCTCATGAGGAATGAAGAATAAACACATTAAAGTCTTACCATTTGCTAAATTAA
 GTCTGGCTCTTGTACCAAATTCTGCTTGTGCTGTAAATTATATTGTATATT
 TTCTATCAACATTTTACTGTGTGGTGTGTTGTAATTAAACGTTAAAGCAAAC
 TCAGAACAAATGAATTCTCACGAATATTCACTGATATTACAGTTGAGAAATAACTACTTC
 TGTAGTAGGTAATTAAATGTCCAATGCAAGTTAACGTGTCACTGATCACGCTATTCA
 [G, A]
 GTGTGTGTTGATAAGGGGAGGTGGGGAGTTGTGGGTTGATTTTATTGCTTTC
 TCATGTGACTGTGTGTCATGTTAGTAAACAAATGGTTGCGAGAGAACAGTAGTCTTTG
 CAAAGATTGCTTATACAGAGCACTCAATTCTCATATTATTATAATGGCTTAATTAA
 AGCCTTAAATTATTAGAAACTCATAAATAATTGTTATTGTTTGTGAGATGGAGTT
 TCGCCCTTATTGTCAGGCTGAAGTACAATGATGTGATCTGACTCACTGCAACCTCCGC

27253 GCTTAAGCCATGCATGGGTTTATAGGAGATGTAGTCTCACAGTGAGTTATTGTA
 GCTGTGTTTGTGTTGTTAGCTTATAGCAATGCAAGTGTGCTTTATTAAACATCATT
 TTCTTTCTTTGCAGTGATTATTCAAGTACTCTGATTGGCGACTCAGGGGT
 TGGAAAGTCTTGCCTCTCTAGGTTGCAAGTTGAAATTGAAATGTCTTACAAT
 TAATGGTACAATTAAATGCTATGTATGTTCTAGGTAGATAAAATTAAACAGTTTATTCA
 [A, C]
 GAATAAGTTAATTCTCCAGAATTATATTAAAGACTCCAAATATACATCCCCAGTG
 GTATCTGGACTGTTAAATAGAAAAATTGTTGCTCTAAAGAAATTCACTGAAAGTCT
 GGTTATAAAAGTCAGAATGCTAATACTTTGGTCAGGTCAAACAGCAGTCCAATATAG
 GCAGCAAGTTAAGGGTACTTGGTGGCCTGTGTTGAAAGCGACTTGATGAAATAATC
 TTTAAATTAAACTTAGTAGAATAAAAAGAAAAGCAGAGCCAGGTGACGCACTG

28098 CTTTAAATTAGCATGTTCTGGCCAGGTGCGGTGGCTCACGCCGTAACTCCAGCACT
 TTGGGAGGCCGAGACGGGCGATCACAGGTCAAGGAGATTGAGACCATCCTGGCTAACAC
 GGTGAAACCCCGTCTACTAAAAATACAAAAATCAGCTGGGTGTGGTGCCACACGCC
 GTAGTCCCAGCTACTCGGGAGGCTGAGGCAAGGAGAACGCTTGAACCCAGGAGGCCAGG
 TTGCACTGAGCTGAGATGGTGCACACTGCACTCCAGCCTGGCAACAGAGCAAGACTGTCTC

[-, A]

AAAAAAAAAGAAAAAAATAAAAACAAATTAGCATGTTCCCTTAGAGATCATTGT
 TTCTCAGAGCATGGACAAAGACTCCTGGGGTTACCAAGACCCCTCAGTAGGCCATG
 AGGTCAAATATCCTAATAACTAAGATGTTAGTATTTGTAAGGAAATATTACTGGT
 AATAACTAATATAAAAGATGTTGCCCTTTCAGTGATGACATTGGCTCTGGTACAAA
 AGCATGTGGGTTAAATTGCTGCTGGCTGGTACACATCAAGGCAGCGCTAACGCTCCAAAT

28597 GATGTTTGCCTTTTCAGTGATGACATTGGCTCTGGTACAAAAGCATGTGGTAAAATTG
 CTGCTGGCTTGGTACACATCAAGGCAGCGCTAACGCTCCAAATTGACTCATGGTGATGCC
 ATTCTTACCTCTGTGCCCTCACAGGAACAAAACAAGCGTGCCTTTTATTGAAGAT
 TGTCCCTGACAAAACAGTTAAATGATTAATTGGAAATGTTGATCCATGAGTATTC
 CTTTAAAATATTGTGAAGAAATGGGAAGTTACATAAAACAATGTTTTTTGTTT
 [G, T]
 TTTTTTTTTTTTTGAGACAGATTCTGGCTGTGTTGCCAAGGCTAGAGTGCAGTGGC
 GTCTGGCTCCCAGGCTCAAGCTGTTCTCCACTTCAGCCTCCAAAGTGGCTGGACCTCC
 CAAGTGGATGCCCATCATGCCCTGGTGAATTGTTGTTGAGTGCACAGGTCTC
 ACTGTGTTGCACAGGCTGGTCTAAACCTCTGAGCTAACGCGATGCATGTGCCCTCAGCCT
 CCCAAAGTGCAGGAAAGCACTTTTACTGCATACTGGTAGTGTGGTTATTG

31431 CTGCATTTTTTTTTTTGGTTGAGATGGAGTCTCGCTCTGCGCCAGGCTGGA
 GTGCAGTCGTGCAATCTCGGCTCACTGCAGCCTCCACCTCATGGGTTCAAGCGATTCTCC
 ATCTGGTCTCTGACTAGCTAGGTTACAGGCCTGCCCCACTAACCCACTAATT
 GTATTTTAGTAGAGACAGGGTTTACCATGTTGCCAGGCTGGTCTGAACCTCTGATC
 TAAAGTGAGCCTCCACCTGGCCTCCAAAGTGCTGGATTACATATGTGAGCCACTGC
 [C, T, G]
 CCTGGCCTCTATATACTTCTATAGTACCTGATACTTATTAGGCACCAATTACAACATAA
 CTTTTTTTTTTTTTTTGAGACAGAGACATGCCCTGCGCTGGGCTGGAGTGC
 AGTGGCACAGTCTCGGCTCACTGCAACCTCACCTCCGGGTTCAAGTGATTCTCCTTCC
 TCAGCCTCCGGTAGCTGGATTACAGGCCTGCCACCACGCTCAGCTAAATTGTTG
 ATTTTAATAGAGATGAGGTTTACCATCTGGCCAGGCTGATCTCAAACCTCTGACCTT

35704 ATGTGTGATCATTGGTGTATAAGATTTGGGTGTATTGCGTGTGAAACATTCAAT
 TTTGTTACTTCTGTGGCTGGAGGGATCTTATAGGACACTGTCCTTCATCTTGTC
 TCTTCATCTTAAAGGAATTCTTTCATGCCCTGAAGGCTCATTTGAAACATT
 TTTGTTGTTTTTATTTTGAGATACAGTATTGCTCTGTCCTCCAGGCTGGAGTGCA
 GTGGCGCGATTGAGCTCACTGCAACCTCCGCCCTGGTTCAAGTGATTCTCCTGCCT
 [C, T]
 AGCCTCCCTAATAGCTGGGATTACATGTTGTAACACCATGCCCGACAATT
 TTTGAGATGGAGCCTTGTGTTGCGCCAGGCTGGAGTGCAGTGGTGAATCTGGCTC
 GCTGCAGCCTCCGCCCTCCAGGTTCAAGCAGTCTCTGCTCAGCCTCCTGAGTAGCTG
 GGATTACAGGCGTGCCTGCCACCACCCCTGCTAATTGTTGTTAGTAGAGACAGAG
 TTTCACCATGTTGGTAGGCTGGTCTCGAACCTCGTGTACTGCCTGACTCGC

35728 GATTGGGTGTATTGCGTGTGAAACATTCAATTGTTACTTCTGTGGCTGGAA
 GGGATCTTATAGGACACTGTCCTTCATCTTGCTGTCTTCATCTTAATAGGAATT
 TTTTCATGCCCTGAAGGCTCATTTGAAACATTGTTGTTGTTTATTGAA
 GATACAGTATTGCTCTGTCCTCCAGGCTGGAGTGCAGTGGCGCAATTGAGCTCACTGCA
 ACCTCCGCCCTCTGGGTTCAAGTGATTCTCCTGCCCTAGCCTCCCTAATAGCTGGGATT
 [C, T]
 ATGTGTGTAACACCATGCCCGACAATTGTTGAGATGGAGCCTGCTTGTC
 GCCCAGGCTGGAGTGCCTGGTGAATCTTGGCTCGCTGCAGCCTCCGCCCTCAGGTT
 CAAGCAGTCTCTGCTCAGCCTCCTGAGTAGCTGGATTACAGGCGTGCCTGCCACCACA
 CCCTGCTAATTGTTGTTAGTAGAGACAGAGTTCACCATGTTGGTAGGCTGGT
 CTGAAACTCCTGACCTCGTGTACTGCCTGACTCGGCTCCAAAGTGCTGGATTACAGG

36690 AAAAAAAAGTAACCAGGTGGTGGTCCATGCCCTGAGTCCTAGCTCCCCAG
 GAGACTGAGGTGGAGGAATGTTGAGGCCAGGACTCAAGGCTGCAGTGAGGCAAGATT
 GCACCATTCACCCAGTTGGGACAGAGTGAGAGACCCCTGCTCAAAAACAAAATAA
 GGCTGGCGCAGTGGCTGCCGGCTGTTCAAGCTTATAGTCCTAGCAGCCTGG
 AGGCCAAGGTGGCAGATTGCCCTGAGCTCAGGAGGCTAACGACCAAGCCTGAGCAACATGG
 [C, T]
 GAAACCTCATCTTGCAAAACATACAGAAAAACAAAAACACAAACCTCTAGTT
 GCCAGTTATTGTTATTCTAGTGATTCTCTTTCTTTCTGAGACAA

AAATTCACTTGTCTCCCTCGCTAGAGTGCAGCGGTAGCTCACTACATGATTCTTTA
 GAGACATGTTAATTCTTATATTGAGCTGAAGCCTGTTCTTACTTCTGCTCTTCTT
 ATTCCCTCCGCCCTGTAGAGCTGCCTGAATCAGATTAATTCTCTTTATTGGCAAGCCTG

41002 GAGTTGAGGACTAATGTTCTATATCACATCCTGATAATCTCACAGTTATGAAAATCAA
 ACTATTCCCCTCCCTCACACTTTCCCCAACCTTATTAAATGGAATTGTTGGATT
 TCTGATTGTTGTAATAGTGGACACAGCAGGCCAGGAAAGATTCGAACAATCACCT
 CCAGTTATTACAGAGGAGGCCATGGCATCATAGTTGTATGATGTGACAGATCAGGAA
 GTTCCAAGAGGAGATTGTGTTACAGTGACCAAGTAGGAAGCCATTATTGATTAATGTCA
 [G, C]
 ATTCATTTACTACTTCATATATAAGCCATCAGTATTAATTAAATGGCAGAAAACCTTGTC
 CACTCTCAAATATAAATGTGAATCACTAAAAGACATTGTTCTGTAATAAAATAAAA
 GATTAGTAATTAGTTTACGTTGCTTCAAGGGATTCTGGTTGATTATTGTCAACTA
 AATAACTTGATCAAATAGCCAAGACTCTAACATATAGGCAAGAGTTGTAGGGAAATCGT
 GAGTTGCTTGGCTTAACTGTGTTCTGGTGTAAAGTATTAACAGGAATATGGCCTGGT

41033 CTGATAATCTCACAGTTATGAAAATCAAATATTCCCCTCCCTCACACTTTCCCC
 AACTTTATTAAATGGAATTGTTGGATTCTGATTGTTGTAATAGTGGACACAGC
 AGGCCAGGAAAGATTCGAACAATCACCTCCAGTTATTACAGAGGAGGCCATGGCATCAT
 AGTTGTGATGATGTGACAGATCAGTAAGTCCAAGAGGAGATTGTGTTACAGTGACCA
 AGTAGGAAGCCATTATTGATTAATGTCAAGATTACTTCACTATATAAGCCATC
 [A, G]
 GTATTAATTAAATGGAAGAAAATTTGTCACACTCAAATATAAATGTGAATCACTTAA
 AGACATTGTTCTGTAATAAAATAAAAGATTAGTAATTAGTTACGTTGCTTCAA
 GGGATTCTGGTTGATTATTGTCAACTAATAACTTGTCAATAGCCAAGACTCTAA
 CATATAGGCAAGAGTTGTAGGGAAATCGTGAGTTGCTGGTTAACTGTGTTCTGGTG
 TTAAGTATTAACAGGAATATGGCCTGGTAATTAGAACATTGTCCATCAGAATTGCCAAAAG

43161 AGTCCTTCAATAATGTTAACAGTGGCTGCAGGAAATAGATGTTATGCCAGTGAAAATG
 TCAACAAATTGTTGGTAGGAAACAAATGTGATCTGACCACAAAGAAAGTAGTAGACTACA
 CAACAGCGAAGGTATTTAAAGTTATTTCATACTGAATTGAAAGTGTGAATTAT
 GTATGGGTTCTGCAAGTAAAGGCCACAGCCTTTAAAATATGTGCACTAGAACATAC
 TGTGACAGTGACAATTGTTGTTAGCATCTGTTGGATCCAATGAACCTAGTCCCTCACGCT
 [C, T]
 CATTATGGATGGTAGAAATGCAAGTAAGAATTAGTGGAAAAGATTTCAGTGTAAATTGT
 GCCTCATTATTCTCTTAGGAATTGCTGATTCCTTGGAAATTCCGTTTGAAACCAAGT
 GCTAAGAATGCAACGAATGTAGAACAGTCCTCATGACGATGGCAGCTGAGATTAAAAG
 CGAATGGGTCGGAGCAACAGCTGGTGTGAGAAGTCCAATGTTAAAATTCAGAGC
 ACTCCAGTCAGCAGTGGAGGTGCTGTTAAATTGCTCCATCCTTTCTCAC

43765 AATGAATTGCAATCTGAACCCAAGTGAAAAACAAAATTGCCGTGAAATTGTAAGTGTATGT
 AGCTGCACTAACACAGATTCTTACCGTCTCCACAAAGGTAGAGATTGTAATGGTCAAT
 ACTGACTTTTTTTATTCCCTTGACTCAAGACAGCTAACCTCATTTCAGAACTGTTT
 AACACCTTTGTGCTGGTTATAAAATGTGTTAATCCTGTTGCTTCCGTGATACC
 AGACTGTTCCCGTGGTTGAGAATATATTGTTGATGTTATATTGGCATGTT
 [A, G]
 GATGTCAGGTTAGTCTCTGAAAGATGAGTTGAGCCATTGTTGATCAAACAGCACAAGC
 AGTGTCTGTCATTCCATGCTAAAGTTAGTGGAGATGTTATGTAAGATCTGATTG
 CTAGTTCTCCTGCTAGAGTTATAATGGAAAGATTACACTATCTGATTAATAGTTCTT
 CATACTCTGCATATAATTGCTGGCTGAGAATATTGTAATTGTCACACTATGTAACA
 AAACAACTGAAGATATGTTAATAAAATTGTACTTATTGGAAGTAATATCAAACGTAT

44713 AACAGCACCTTCTAATTGGCAAATGATCAGACTAATGTTGCTAATGTTTCTTCC
 ATGTTTCAGTCAGATTCAACTATTAAATCCTCCACAGTTGCTTAACCTGGTGTGGAGG
 AGGGTTAAGCATTAAGATAGGAAGCAGGAAATTGATGTTGCTCTAAATTAGAAATTATA
 TCCCTAAAAATTAAACATGATAACTGGTGGTAATGATAATTGAGGCAAATGTATTAT
 TTGGTGACATTGCTGATATGAAGATTCTGAAATAGGACCTCAAGATCCTAGGGG
 [G, T]
 TTTTGGTTGGTTTAATTGTGAGGAATAAAAATCTCTGCCACACTGGCATTAAAG
 GTGACTGAGGTCAAACGTTGTTCTTAGGTTGAAATAGCAGCCAAAATCTCACGC
 AGGGGCTTGGGATATGGCTGGCAACACATTGTTGTTGGCTCCTTAATTAAATGAT
 AAAATTAAAGCTAACACAGCCAAAATGAATAGGTTTTTTAATTGTTATTTTCACT
 AACACAGGAATTGAAATACATGGTACAAAATAAGTGTAAAGATAATTGTAATTGAAAT

44831

GGAGGGTTAACATTAAGATAGGAAGCAGGAAATTGATTGCTCTAAATTAGAAATTA
TATCCCTAAAAATTAAAACATGAATACTGGGTGGTAATGATAATTGAGGCAAATGTATT
ATTTGGTGACATTTGCATATATGAAGATTITCTGAAATAGGACCTCAAGATCCTAGG
GGGTTTGTGGTTTAATTGTGAGGAATAAAAATCTCTGCCACACTGGCATTTT
AAGGTGACTGAGGTCAAACGTTGTTCTTAGGTTGAAATAGCAGCCAACATTCTTC
[C, T]
GCAGGGGCTTGGGATATGGCTGCTGGCAACACATTTGTTGTGGCTCTTAATTAAATG
ATAAAATTAAAGCTAACACAAGCCAAAATGAATAGGTTTTTAATTTTTATTTC
CTAACACAGGCAATTGAAATACATGGTACAAAATAAGTGGTAAGATAATTGTAAAATGAA
ATGGACAGAAATTCAATTTCATCTATGAAAATTTCACAATAAAAATCATAGTTACT
TTGTATTATAGCGTGCTGGTGGATCTATTCATCCTCACATAAGGCAACTGACAAATTCA